81844

DECLARATION

I, Jane Roberta Mann, B.A., a Translator, of Frank B. Dehn & Co., 59 St Aldates, Oxford OX1 1ST, England, do declare that I have a competent knowledge of the English and German languages and that the document that is annexed hereto is a true and accurate translation of the German text of the U.S. Provisional Application Serial No. 60/431,535 filed December 06, 2002.

I further declare that all statements made of my own knowledge are true and that all statements made on information and belief are believed to be true.

I acknowledge that wilful false statements and the like are punishable by fine or imprisonment, or both [18 U.S.C. 1001] and may jeopardize the validity of the application or any patent issuing therefrom.

Jane Mann

Signed this 5th day of November, 2003

Eukaryotic expression vector comprising a modified neomycin-phosphotransferase gene, and methods for the selection of recombinant cells producing high levels of a desired gene product

Scope of the Invention

The invention relates to a method for selecting highly productive recombinant cells, a method for preparing heterologous gene products and expression vectors which contain a modified neomycin-phosphotransferase-gene and a gene of interest functionally linked to a heterologous promoter and host cells transfected therewith which may be used in these processes.

Background to the Invention

Mammalian cells are the preferred host cells for the production of complex biopharmaceutical proteins as the modifications carried out post-translationally are compatible with humans both functionally and pharmacokinetically. The main relevant cell types are hybridoma, myeloma CHO (Chinese Hamster Ovary) cells and BHK (Baby Hamster Kidney) cells. The cultivation of the host cells is increasingly carried out under serum- and protein-free production conditions. The reasons for these are the concomitant cost reduction, the reduced interference in the purification of the recombinant protein and the reduction in the potential for the introduction of pathogens (e.g. prions and viruses). The use of CHO cells as host cells is becoming more widespread as these cells adapt to suspension growth in serum- and protein-free medium and are also regarded and accepted as safe production cells by the regulatory authorities.

In order to produce a stable mammalian cell line which expresses a heterologous gene of interest (GOI), the heterologous gene is generally inserted in the desired cell line together with a selectable marker gene such as e.g. neomycin phosphotransferase (NPT) by transfection. The

heterologous gene and the selectable marker gene can be expressed in a host cell starting from one individual or separate cotransfected vectors. Two to three days after transfection the transfected cells are transferred into medium containing a selective agent, e.g. G418 when using neomycin phosphotransferase-gene (NPT gene), and cultivated for some weeks under these selective conditions. The high-growing resistance cells which have integrated the exogenous DNA can be isolated and investigated for expression of the desired gene product (of the GOI).

A major problem in establishing cell lines with a high expression of the desired proteins arises from the random and undirected integration of the recombinant vector into transcriptionally -active or -inactive loci in the host cell genome. As a result a population of cells is obtained which have completely different expression rates of the heterologous gene, the productivity of the cells generally following normal distribution. In order to identify cell clones which have a very high expression of the heterologous gene of interest it is therefore necessary to examine and test a large number of clones, which is time consuming, labour intensive and expensive. Improvements to the vector system used for transfection therefore set out to increase the proportion of high producers in the transfected cell population by suitable selection strategies and thereby reduce the expenditure and work involved in clone identification. The development of such an expression system is the subject of the present invention.

The amino glycoside-3'-phosphotransferase II enzyme (neomycin-phosphotransferase) (EC27195) the gene of which is 5- associated in *Escherichia coli* transposons is used as a selectable marker in a number of organisms (e.g. bacteria, yeasts, plants and mammalian cells). This enzyme confers resistance to various aminoglycoside antibiotics such as neomycin, kanamycin and G418, by inactivating the antibiotics by transferring the terminal phosphate from ATP to the 3' hydroxyl group of the aminohexose ring I. In addition to the wild-type neomycin phosphotransferase some mutants are known which have reduced phosphotransferase activity and hence reduced resistance to aminoglycoside antibiotics in bacteria (Blázquez et al.,

1991; Kocabiyik et al., 1992; Yenofsky et al., 1990) and in slices of leaf from tobacco (Yenofsky et al., 1990).

One of these mutants (Glu182Asp) was used as a marker for selecting embryonic stem cells, the neomycin phosphotransferase gene being integrated into the c-myc gene by targeted homologous recombination (gene targeting) (Hanson et al., 1995). The authors restrict themselves to the use of the modified enzyme for gene targeting.

Patent application WO 99/53046 describes the expression of a modified neomycin phosphotransferase gene (Asp261Asn) in production-relevant mammalian cells. The authors describe a non-cloning method for expression of a gene of interest in mammalian cells. By cotransfection of the cells with three individual DNA fragments which code for a promoter element, a gene of interest and a selectable marker coupled with an IRES ("Internal ribosomal entry site") element, it is possible to deliberately grow cells, under selection pressure, in which all three DNA fragments are combined as a functional bicistronic transcription unit (promoter gene of interest-IRES-neomycin-phosphotransferase gene). The arrangement of the elements only occurs in the transfected cell, so that only a few cells show the correct arrangement of the elements. Moreover, after gene amplification, using an amplifiable selectable marker, no high producing clones can be generated. After repeated selection and gene amplification the cells generated exhibited at most 6 pg of protein per cell per day (6pg/cell/day).

None of the publications provides a high expression vector system for developing high producing cells in order to prepare recombinant biopharmaceutical proteins in which one or more complete functional transcription units code or codes both for one or more genes of interest and also for a mutated neomycin phosphotransferase gene with reduced antibiotic resistance. The DNA construct described in WO 99/53046 contains only a promoter-less neomycin gene functionally linked to the gene for dihydrofolate reductase (DHFR).

There is therefore a need to make such high expression vector systems available for pharmaceutical processes. The problem of the present invention was therefore to provide a method of selection for high producing recombinant cells, a process for producing heterologous gene products and expression vectors which contain a modified neomycin phosphotransferase gene and a gene of interest functionally linked to a heterologous promoter, which will enable high producing cell clones to be selected.

In this context the problem was further to identify suitable modified neomycin-phosphotransferase genes which are particularly suitable for selecting biopharmaceutically relevant, high producing cell clones.

Surprisingly, it has been found that high producing cell clones which exhibit substantially higher expression rates than those described in WO 99/53046 can be produced with the vector configuration according to the invention described here. This is particularly true when using selected neomycin-phosphotransferase genes.

Summary of the Invention

Four different modified neomycin-phosphotransferase genes (Glu182Asp, Asp190Gly, Asp208Gly, Asp227Gly described in Yenofsky et al., 1990; Kocabiyik et al., 1992; Hanson et al., 1995) in stable transfections of CHO-DG44 cells (Urlaub et., 1983) were tested for their suitability as marker genes for selectively enriching high producing cells, by comparison with the wild-type gene.

Surprisingly, it was found that when the neomycin phosphotransferase mutants Glu182Asp and Asp227Gly were used as selectable markers an enrichment of transfected mammalian cells with high expression rates for the cointegrated gene of interest could be achieved, whereas the Asp190Gly and Asp208Gly mutants proved to be unsuitable markers for selecting transfected CHO-DG44 cells under serum-free culture conditions.

By comparison with the use of wild-type neomycin phosphotransferase as selectable marker, the cells exhibited a productivity increased by a factor 1.5 to 2.4 in the case of Glu182Asp mutant and a productivity increased by a factor of 1.6 to 4.1 in the case of the Asp227Gly mutant. These results were obtained both in the expression of a single-chained protein and also in the expression of a multi-chain protein, an antibody.
The problems on which the invention is based were solved using a eukaryotic expression vector which contains a heterologous gene of interest functionally linked to a heterologous promoter and a modified neomycin phosphotransferase gene which codes for a neomycin phosphotransferase which has low enzyme activity by comparison with the wild-type neomycin phosphotransferase.

The modified neomycin phosphotransferase gene is preferably a mutant which codes for a different amino acid from the wild-type gene at amino acid position 182 or 227. In a particularly preferred embodiment the neomycin phosphotransferase gene is the mutant Glu182Asp or Asp227Gly.

The expression vector preferably contains other regulatory elements, e.g. one or more enhancers functionally linked to the promoter or promoters.
Expression vectors are also preferred which additionally contain a gene for a fluorescent protein which is functionally linked to the gene of interest and the heterologous promoter, preferably via an internal ribosomal entry site (IRES), which enables bicistronic expression of the gene which codes for a fluorescent protein and of the gene which codes for a protein/product of interest, under the control of the heterologous promoter. Particularly suitable are expression vectors in which the heterologous gene of interest is under the control of the ubiquitin/S27a promoter.

The invention also relates to expression vectors which instead of the gene of interest contain a multiple cloning site for incorporating such a gene, i.e. a sequence section with multiple recognition sequences for restriction endonucleases.

In another aspect the invention relates to recombinant mammalian cells which have been transfected with one of the expression vectors according to the invention. These are preferably recombinant rodent cells, of which recombinant hamster cells such as e.g. CHO cells or BHK cells are particularly preferred. In another preferred embodiment the recombinant cells are additionally transfected with the gene for an amplifiable selectable marker, e.g. with the gene of dihydrofolate reductase (DHFR).

The invention also relates to a process for enriching recombinant mammalian cells which express a modified neomycin phosphotransferase gene, characterised in that (i) a pool of mammalian cells is transfected with a gene for a modified neomycin phosphotransferase, which has only 10 to 50% of the activity compared with wild-type neomycin phosphotransferase; (ii) the mammalian cells are cultivated under conditions which allow expression of the modified neomycin phosphotransferase gene; and (iii) the mammalian cells are cultivated in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells, and gives preference to the growth of those cells which express the neomycin phosphotransferase gene.

The invention also relates to a process for the expression of at least one gene of interest in recombinant mammalian cells, characterised in that (i) a pool of mammalian cells is transfected with at least one gene of interest and one gene for a modified neomycin phosphotransferase which exhibits only 10 to 50% of the activity compared with wild-type neomycin phosphotransferase; (ii) the cells are cultivated under conditions which allow expression of the gene or genes of interest and the modified neomycin phosphotransferase gene; (iii) the mammalian cells are cultivated in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells and gives preference to the growth of those cells which express the neomycin phosphotransferase gene; and (iv) the protein or proteins of interest is or are obtained from the mammalian cells or from the culture supernatant.

The modified neomycin phosphotransferase gene is preferably a neomycin phosphotransferase gene which has been mutated at amino acid position 182

or 227 compared with the wild-type gene, and the modified neomycin phosphotransferase gene preferably codes at amino acid position 182 for aspartic acid or at amino acid position 227 for a glycine. The invention particularly relates to corresponding methods by which the host cells are transfected with one of the expression vectors according to the invention as described.

Accordingly, the present invention relates to a process for obtaining and selecting recombinant mammalian cells which express at least one heterologous gene of interest, which is characterised in that (i) recombinant mammalian cells are transfected with an expression vector according to the invention which in addition to the gene of interest and the modified neomycin phosphotransferase gene codes for a fluorescent protein; (ii) the mammalian cells are cultivated under conditions which allow expression of the gene or genes of interest, the gene which codes for a fluorescent protein and the modified neomycin phosphotransferase gene; (iii) the mammalian cells are cultivated in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells and gives preference to the growth of those cells which express the neomycin phosphotransferase gene; and (iv) the mammalian cells are sorted by flow-cytometric analysis.

If the mammalian cells have additionally been transfected with a gene for an amplifiable selectable marker gene, e.g. the DHFR gene, it is possible to cultivate the mammalian cells under conditions in which the amplifiable selectable marker gene is also expressed, and to add to the culture medium a selecting agent which results in amplification of the amplifiable selectable marker gene.

Preferably, the processes according to the invention are carried out with mammalian cells which are adapted to growth in suspension, i.e. with mammalian cells which are cultivated in a suspension culture. Other embodiments relate to processes in which the mammalian cells, preferably those which are adapted to growth in suspension, are cultivated under serum-free conditions.

Description of the Figures

Figure 1 shows a diagrammatic representation of the base vectors used to express the recombinant proteins in CHO-DG44 cells. "P/E" is a combination of CMV enhancer and hamster-ubiquitin/S27a promoter, "P" on its own indicates a promoter element and "T" is a termination signal for transcription, which is needed for the polyadenylation of the transcribed mRNA. The position and direction of transcription initiation within each transcription unit is indicated by an arrow. For cloning the heterologous genes a sequence region with multiple cutting sites for restriction endonucleases (multiple cloning sites - MCS) is inserted after the promoter element. The amplifiable selectable marker dihydrofolate reductase is abbreviated to "dhfr" and the selectable marker neomycin phosphotransferase is abbreviated to "npt" or "npt" D227G (Neomycin Phosphostransferase Mutant). The "IRES" element coming from the encephalomyocarditic virus acts an internal ribosomal entry site within the bicistronic transcription unit and enables translation of the following green fluorescent protein "GFP".

Figure 2 shows a diagrammatic view of the eukaryotic expression vectors which code for a biopharmaceutical protein and are used to transfect CHO-DG44 cells. "P/E" is a combination of CMV enhancer and hamster ubiquitin/S27a promoter, "P" on its own is a promoter element and "T" is a termination signal for the transcription which is needed for the polyadenylation of the transcribed mRNA. The position and direction of transcription initiation within each transcription unit is indicated by an arrow. The amplifiable selectable marker dihydrofolate reductase is abbreviated to "dhfr" and the selectable marker neomycin phosphotransferase is abbreviated to "npt". The IRES element originating from the encephalomycoarditis virus acts as an internal ribosomal entry site within the bicistronic transcription unit and permits translation of the following green fluorescent protein "GFP". MCP-1 codes for monocyte chemoattractant protein-1, whereas "HC" and "LC" code for the heavy and light chains, respectively, of a humanised monoclonal IgG2 antibody.

Figure 3 shows the part of the sequence of the neomycin phosphotransferase (npt) gene in which the point mutations have been inserted by PCR with mutagenic primers. The capital letters indicate the nucleotide sequence of the npt coding region whereas the small letters indicate the flanking non-coding nucleotide sequences. The amino acid sequence predicted from the nucleotide sequence (3-letter code) is shown above the coding nucleotide sequence. Arrows indicate the direction, length and position of the primers used, the arrows with solid lines indicating the mutagenic forward primers, the broken lines indicating the mutagenic reverse primers, the dotted lines indicating the primer Neofor5 located upstream of the npt gene (SEQ ID NO:11) and the dot-dash line indicating the primer Neorev5 located downstream of the npt gene (SEQ ID NO:20). The nucleotides exchanged with respect to the wild-type sequence are emphasised above and below the arrows.

Figure 4 is a diagrammatic view of the eukaryotic expression vectors each of which codes for a biopharmaceutical protein and contains as selectable marker a mutated neomycin phosphotransferase (npt). The NPT mutants E182D (SEQ ID NO:3), D190G (SEQ ID NO:5), D208G (SEQ ID NO:7) and D227G (SEQ ID NO:9) contain a point mutation which results in a modified amino acid (given in 1-letter code) at the position indicated. P/E is a combination of CMV enhancer and hamster ubiquitin/S27a promoter, while P is just one promoter element and T is a termination signal for transcription which is required for the polyadenylation of the transcribed mRNA. The position and direction of the initiation of transcription within each transcription unit is indicated by an arrow. MCP1 codes for the monocyte chemoattractant protein 1 whereas LC codes for the light chain of a humanised monoclonal IgG2 antibody. The IRES element originating from the encephalomyocarditis virus acts as an internal ribosomal entry site within the bicistronic transcription unit and permits translation of the following green fluorescent protein GFP.

Figure 5 shows the influence of the npt mutations on the selection of stably transfected MCP-1 expressing cells. For this, CHO-DG44 cells were

transfected with the vectors pBIN-MCP1, pKS-N5-MCP1 and pKS-N8-MCP1 (Figs. 2 and 4), which contained as selectable markers either the npt wild-type (WT) or the npt mutants Glu182Asp and Asp227Gly. In order to select stably transfected cells, 400 µg/mL or 800 µg/mL of G418 were added to the medium as selective agent. The concentration of the recombinant proteins MCP-1 produced in the cell culture supernatant was measured by ELISA and the specific productivity per cell and per day was calculated. Each bar indicates the mean of the specific productivity of the titre from 18 pools over four cultivation runs in 6-well dishes.

Figure 6 shows the influence of NPT mutations on the selection of stably transfected mAb expressing cells. For this CHO-DG44 cells were transfected with the plasmid combinations pBIDG-HC/pBIN-LC (NPT wild-type), pBIDG-HC/pKS-N5-LC (NPT mutant Glu182Asp) or pBIDG-HC/pKS-N8-LC (NPT mutant Asp227Gly) which differ from another only in the NPT gene used as selectable marker (wild-type or mutant). The concentration of the recombinant monoclonal IgG2 antibody produced in the cell cultures supernatant was determined by ELISA and the specific productivity was calculated per cell and per day. In all, four independent transfection series (T1 to T4) were carried out, each transfection series consisting of 6 cell pools. The bars represent the averages of the specific productivity or the titre from the 6 cell pools of a transfection series of 6 to 8 cultivation runs in 75cm² flasks.

Figure 7 shows another transfection series in which CHO-DG44 cells are transfected with the plasmid combinations pBIDGHC/pBIN-LC (NPT wild-type), pBIDG-HC/pKS-N5-LC (NPT mutant Glu182Asp) or pBIDG-HC/pKS-N8-LC (NPT mutant Asp227Gly), which differ only in the NPT gene (wild-type or mutant) used as the selectable marker. The concentration of the recombinant monoclonal IgG2 antibody produced in the cell culture supernatant was determined by ELISA and the specific productivity was calculated per cell and per day. Each bar indicates the mean specific productivity or the titre from 9 pools over 6 cultivation runs in 75 cm² flask.

Figure 8 shows the concentration of cells with a higher GFP expression in transfected cell pools by using an NPT mutant as selectable marker. For this, CHO-DG44 cells were transfected with the plasmid combinations pBIDGHC/pBIN-LC (NPT wild-type), pBIDG-HC/pKS-N5-LC (NPT mutant Glu182Asp) or pBIDG-HC/pKS-N8-LC (NPT mutant Asp227Gly), which differ from one another only in the NPT gene (wild-type or mutant) used as the selectable marker. Moreover, the pBIDG vectors also contained the GFP as marker gene. After 2 to 3 weeks' selection of the transfected cell pools in HT-free medium with the addition of G418, the GFP fluorescence was measured by FACS analysis. Every graph, with the exception of the non-transfected CHO-DG44 cells used as a negative control, represents the average GFP fluorescence from 9 pools.

Figure 9 shows the enzyme activity of the NPT mutants compared with the NPT wild-type in a dot assay. For this, cell extracts were prepared from two different cell pools (pool 1 and 2) expressing mAb, which had been transfected and selected with the NPT wild-type gene or with the NPT mutants Glu182Asp or Asp227Gly. Non-transfected CHO-DG44 cells were used as negative control. G418 was used as the substrate in the phosphorylation assay. The extracts were filtered through a sandwich of P81 phosphocellulose and nitrocellulose membrane in a 96 well vacuum manifold. Proteins phosphorylated by protein kinases and also non-phosphorylated proteins bind to the nitrocellulose, whereas phosphorylated and non-phosphorylated G418 passes through the nitrocellulose and binds to the phosphocellulose. The radioactive signals were detected and quantified using a phosphoimager.

Figure 10 shows the Northern Blot analysis of NPT expression in the transfected cell pools. For this, total RNA was prepared from two different cell pools expressing mAb (pools 1 and 2) which were transfected and selected either with the NPT wild-type gene or with the NPT mutants Glu182Asp or Asp227Gly. Untransfected CHO-DG44 cells were used as the negative control. 30 μg of RNA was hybridised with a FITC-dUTP-labelled PCR product which comprised the coding region of the NPT gene. The exposure

time of the X-ray film was 1 hour (a) or 24 hours (b). In all the transfected cells a specific singular NPT transcript of about 1.3 kb was detected.

Figure 11 shows the dot blot analysis of transfected cell pools. In this, genomic DNA was prepared from two different cell pools (pool 1 and 2) expressing mAb, which had been transfected and selected either with the NPT wild-type gene or with the NPT mutants Glu182Asp or Asp227Gly. Non-transfected CHO-DG44 cells were used as the negative control. 10 µg, 5 µg, 2.5 µg, 1.25 µg, 0.63 µg and 0.32 µg of genomic DNA was hybridised with a FITC-dUTP-labelled PCR product which comprised the coding region of the NPT gene. The exposure time for the X-ray film was 2 minutes.

Figure 12 summarises the average titres and productivities of cell pools obtained by stable transfection of CHO-DG44 with pBID-HC and pBING-LC. Stably transfected cells were selected in HT-free CHO-S-SFMII medium with 400 µg/mL of G418. After 2 to 3 weeks the concentration of the recombinant monoclonal antibody IgG2 produced in the cell culture supernatant was determined by ELISA and the specific productivity per cell and per day was calculated. Each bar indicates the mean specific productivity or titre from 7 to 8 cultivation runs in 75 cm² flasks.

Figure 13 shows the isolation of high-expression mAb F19 cell pools by a GFP-based selection using FACS with the example of cell pools 5 and 8. These cell pools, obtained by co-transfection with the vectors pBID-HC and pBING-LC, were subjected to sequential GFP-based FACS sorting. The concentration of the IgG2 antibody in the cell culture supernatant of the pools was determined by ELISA after each sorting step and the specific productivity per cell and per day (pg/cell*day) was calculated. A total of 6 sortings were carried out, in which 5% of the cells with the highest GFP fluorescence were sorted out. Each data point represents the average of at least 6 cultivation runs in 75 cm² flask.

Figure 14 shows the increases in mAb productivity achieved by combining GFP-based selection with an MTX amplification step, taking cell pools 5 and 8 as examples. Two weeks after the co-transfection of CHO-DG44 with vectors pBID-HC and pBING-LC the 5% cells with the highest GFP fluorescence were sorted out from pools 5 and 8. Then dhfr-mediated gene amplification was carried out by adding 100 nM methotrexate (MTX) to the culture medium. The concentration of the mAb in the cell culture supernatant of the pools was determined by ELISA and the specific productivity per cell and per day (pg/cell*day) was calculated. Each data point represents the average of at least 6 cultivation runs in 75 cm² flask.

Figure 15 shows the correlation between the antibody productivity and the GFP fluorescence taking cell pools 5 and 8 as examples. These cells were obtained from the transfection of CHO-DG44 with the vector combination pBID-HC and pBING-LC. These pools were subjected to sequential GFP based FACS sorting, sorting out 5% of the cells with the highest GFP fluorescence. After each sorting step the concentration of the IgG2 antibody in the cell culture supernatant of the pool was determined by ELISA and the specific productivity per cell and per day (pg/c*d) was calculated. Each data point represents the average of at least 6 cultivation runs in 75 cm² flasks.

<u>Detailed Description of the Invention and Preferred Embodiments</u>

The present invention relates to processes for preparing and selecting mammalian cell lines which allow high expression of heterologous gene products, preferably biopharmaceutically useful polypeptides or proteins. The processes according to the invention are based primarily on the selection of cells which in addition to the gene of interest express a selectable marker, e.g. neomycin phosphotransferase, which gives the transfected cells a selective advantage over non-transfected cells. Surprisingly, it has been found that the use of modified neomycin phosphotransferase genes (mNPT genes) has a substantial selective advantage over the wild-type neomycin phosphotransferase gene (wtNPT gene). This particularly relates to the use of mutants which have a lower enzyme activity compared with wtNPT.

Neomycin Phosphotransferase genes modified according to the Invention

It has proved particularly suitable to use NPT genes which code for an NPT having only 10 to 50% of the enzyme activity of wtNPT. Preferred NPT mutants are those which have only 20 to 50% of the enzyme activity of wtNPT, while those which have only 25 to 37% of the enzyme activity of wtNPT are particularly preferred. Still further preferred are NPT mutants which have only 25 to 32% of the enzyme activity compared with wtNPT. The enzyme activity of an NPT can be determined for example in a dot assay as described in Example 4 and given as Method 5.

Fundamentally preferred are mutants wherein the reduction in the enzyme activity compared with wtNPT is based on a modification of the amino acid sequence, e.g. on the substitution, insertion or deletion of at least one or more amino acids. Deletion, insertion and substitution mutants can be produced by "site-specific mutagenesis" and/or "PCR-based mutagenesis techniques". Suitable methods are described for example by Lottspeich and Zorbas (1998)(Chapter 36.1) with other references.

The term wild-type neomycin phosphotransferase (wtNPT) refers to a neomycin phosphotransferase gene which codes for the amino glycoside-3'-phosphotransferase II enzyme (EC 2.7.1.95) the gene of which is naturally transposon 5-associated in *Escherichia coli*, and contains for example the amino acid sequence given in SEQ ID NO:2 or is coded by the nucleotide sequence given in SEQ ID NO:1. This enzyme gives resistance to various aminoglycoside antibiotics such as neomycin, kanamycin and G418, by inactivating the antibiotics by the transfer of the terminal phosphate of ATP to the 3' hydyroxyl group of the amino hexose ring I. The term wtNPT also refers to all NPTs which have a comparable enzyme activity to the NPT coded by SEQ ID NO:1. This includes in particular those NPTs in which the enzymatically active centre which catalyses the transfer of a terminal phosphate from ATP to a substrate is present in an identical or nearly identical conformation (Shaw et al., 1993; Hon et al; 1997; Burk et al., 2001) and thus has a comparable enzyme activity to an enzyme which contains the

amino acid sequence of SEQ ID NO:2. A wtNPT has a comparable enzyme activity if it exhibits about 90 to 120% of the enzyme activity displayed by an NPT defined by SEQ ID NO:2, the activity being determined in the dot assay described in Example 4 and referred to as Method 5.

Surprisingly, it has been found that if neomycin phosphotransferase mutants are used as selectable markers in which at least the amino acid glutamic acid at position 182 (Glu182) or aspartic acid at position 227 (Asp227) has been altered compared with wtNPT, it is possible to achieve particularly effective concentration of transfected mammalian cells with a high expression rate for the co-integrated gene of interest. In particular, it has been found that with the mutants Glu182Asp and Asp227Gly as selectable markers it was possible to achieve a concentration of transfected mammalian cells with high expression rates of the co-integrated gene of interest, while the Asp190Gly and Asp208Gly mutants proved to be unsuitable markers for the selection of transfected CHO-DG44 cells under serum-free culture conditions. As a result of the greatly reduced enzyme function of these mutants (Asp190Gly, Asp208Gly) only a few cells were obtained after the selection phase and moreover they had greatly impaired growth and vitality. The amino acids at positions 182 and 227 are non-conserved amino acids which are located outside the three conserved motifs in the C-terminal region of the amino glycoside-3'-phosphotransferases. By contrast the amino acids at positions 190 and 208 belong to the 5 conserved amino acids in the aminoglycoside 3'-phosphotransferases which form the active site (nucleotide binding site) and are located in the first and second conserved motifs (Shaw et al., 1993; Hon et al., 1997; Burk et al., 2001).

Compared with the use of wtNPT as selectable marker the cells in the case of the Glu182Asp mutant showed a productivity increased by a factor of 1.5 to 2.4 and in the case of the Asp227Gly mutant they even showed a productivity increased by a factor of 1.6 to 4.1. These results were obtained both in the expression of a single-chain protein and also in the expression of a multi-chained protein, an antibody. In the former case an expression vector which codes both for the gene of interest and for NPT was used to produce the

stably transfected cells. In the second case, co-transfection was carried out. The two protein chains were each expressed by their own vector, one vector additionally coding for the NPT gene while the other vector coded for the amplifiable selectable dihydrofolate reductase gene.

The present invention thus relates to a process for enriching recombinant mammalian cells which express a modified neomycin phosphotransferase gene, characterised in that (i) a pool of mammalian cells is transfected with a gene for a modified neomycin phosphotransferase which has only 10 to 50%, preferably 20 to 50%, more preferably 25 to 37%, most preferably 25 to 32% of the activity of wild-type neomycin phosphotransferase; (ii) the mammalian cells are cultivated under conditions which allow expression of the modified neomycin phosphotransferase gene; and (iii) the mammalian cells are cultivated in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells, and gives preference to the growth of those cells which express the neomycin phosphotransferase gene.

Particularly preferred is a corresponding process which uses a modified NPT gene described in more detail in this application, particularly if the modified NPT gene used codes for a modified NPT which, by comparison with the wild-type, codes for aspartic acid at amino acid position 182 or glycine at amino acid position 227.

The present invention further relates to eukaryotic expression vectors which contain a heterologous gene of interest functionally linked to a heterologous promoter and a modified neomycin phosphotransferase gene (NPT gene) which codes for a neomycin phosphotransferase (NPT) which has low enzyme activity compared with wild-type neomycin phosphotransferase (wtNPT).

A preferred expression vector contains a modified NPT gene which codes for a modified NPT which has only 10 to 50% of the enzyme activity of wtNPT. Also preferred are expression vectors with modified NPT genes which code for mutants having only 20 to 50% of the enzyme activity of wtNPT.

Particularly preferred are those expression vectors which contain a modified NPT gene which code for mutants having only 25 to 50%, especially 25 to 37% of the enzyme activity of wtNPT. Still more preferred are expression vectors which contain genes of NPT mutants which have only 25 to 32% of the enzyme activity compared with wtNPT.

In another embodiment of the invention the expression vectors contain genes of modified NPT which have been modified, compared with wtNPT, at amino acid position Glu182 or at position Asp227. In this context, NPT mutants are preferred which are modified at position Glu182 or Asp227 and have only 10 to 50%, preferably 20 to 50%, most preferably 25 to 37% of the enzyme activity of wtNPT. Preferably the glutamic acid at position 182 may be replaced by aspartic acid, asparagine, glutamine or another preferably negatively charged amino acid. The aspartic acid at position 227 may be replaced for example by glycine, alanine, valine, leucine or isoleucine.

Particularly preferred are expression vectors which contain modified NPT genes which code for a Glu182Asp or Asp227Gly mutant, which contain the amino acid sequence of SEQ ID NO:4 in the case of the Glu182Asp mutant and amino acid sequence SEQ ID NO:10 in the case of the Asp227Gly mutant. Most preferred is an expression vector using an Asp227Gly mutant, particularly if it has the amino acid sequence given in SEQ ID NO:10 or if it is coded by the nucleic acid sequence given in SEQ ID NO:9.

Gene of Interest

The gene of interest contained in the expression vector according to the invention comprises a nucleotide sequence of any length which codes for a product of interest. The gene product or "product of interest" is generally a protein, polypeptide, peptide or fragment or derivative thereof. However, it may also be RNA or antisense RNA. The gene of interest may be present in its full length, in shortened form, as a fusion gene or as a labelled gene. It may be genomic DNA or preferably cDNA or corresponding fragments of fusions. The gene of interest may be the native gene sequence, or it may be

mutated or otherwise modified. Such modifications include codon optimisations for adapting to a particular host cell and humanisation. The gene of interest may, for example, code for a secreted, cytoplasmic, nuclear-located, membrane-bound or cell surface-bound polypeptide.

The term "nucleotide sequence" or "nucleic acid sequence" indicates an oligonucleotide, nucleotides, polynucleotides and fragments thereof as well as DNA or RNA of genomic or synthetic origin which occur as single or double strands and can represent the coding or non-coding strand of a gene. Nucleic acid sequences may be modified using standard techniques such as site-specific mutagenesis or PCR-mediated mutagenesis (e.g. described in Sambrook et al., 1989 or Ausubel et al., 1994).

By "coding" is meant the property or capacity of a specific sequence of nucleotides in a nucleic acid, for example a gene in a chromosome or an mRNA, to act as a matrix for the synthesis of other polymers and macromolecules such as for example rRNA, tRNA, mRNA, other RNA molecules, cDNA or polypeptides in a biological process. Accordingly, a gene codes for a protein if the desired protein is produced in a cell or another biological system by transcription and subsequent translation of the mRNA. Both the coding strand whose nucleotide sequence is identical to the mRNA sequence and is normally also given in sequence databanks, e.g. EMBL or GenBank, and also the non-coding strand of a gene or cDNA which acts as the matrix for transcription may be referred to as coding for a product or protein. A nucleic acid which codes for a protein also includes nucleic acids which have a different order of nucleotide sequence on the basis of the degenerate genetic code but result in the same amino acid sequence of the protein. Nucleic acid sequences which code for proteins may also contain introns.

The term cDNA denotes deoxyribonucleic acids which are prepared by reverse transcription and synthesis of the second DNA strand from a mRNA or other RNA produced from a gene. If the cDNA is present as a double stranded DNA molecule it contains both a coding and a non-coding strand.

The term intron denotes non-coding nucleotide sequences of any length. They occur naturally in numerous eukaryotic genes and are eliminated from a previously transcribed mRNA precursor by a process known as splicing. This requires precise excision of the intron at the 5' and 3' ends and correct joining of the resulting mRNA ends so as to produce a mature processed mRNA with the correct reading frame for successful protein synthesis. Many of the splice donor and splice acceptor sites involved in this splicing process, i.e. the sequences located directly at the exon-intron or intron-exon interfaces, have been characterised in the by now. For an overview see Ohshima et al., 1987.

Protein/Product of Interest

Proteins/polypeptides with a biopharmaceutical significance include for example antibodies, enzymes, cytokines, lymphokines, adhesion molecules, receptors and the derivatives or fragments thereof, but are not restricted thereto. Generally, all polypeptides which act as agonists or antagonists and/or have therapeutic or diagnostic applications are of value.

The term "polypeptides" is used for amino acid sequences or proteins and refers to polymers of amino acids of any length. This term also includes proteins which have been modified post-translationally by reactions such as glycosylation, phosphorylation, acetylation or protein processing. The structure of the polypeptide may be modified, for example, by substitution, deletion or insertion of amino acids and fusion with other proteins while retaining its biological activity.

Examples of therapeutic proteins are insulin, insulin-like growth factor, human growth hormone (hGH) and other growth factors, tissue plasminogen activator (tPA), erythropoietin (EPO), cytokines, e.g. interleukines (IL) such as IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18 interferon (IFN)-alpha, -beta, -gamma, -omega or -tau, tumour necrosis factor (TNF) such as TNF-alpha, beta or gamma, TRAIL, G-CSF, GM-CSF, M-CSF, MCP-1 and VEGF. Other examples are monoclonal,

polyclonal, multispecific and single chain antibodies and fragments thereof such as for example Fab, Fab', $F(ab')_2$, Fc and Fc' fragments, light (L) and heavy (H) immunoglobulin chains and the constant, variable or hypervariable regions thereof as well as Fv and Fd fragments (Chamov et al., 1999). The antibodies may be of human or non-human origin. Humanised and chimeric antibodies are also possible.

Fab fragments (fragment antigen binding = Fab) consist of the variable regions of both chains which are held together by the adjacent constant regions. They may be produced for example from conventional antibodies by treating with a protease such as papain or by DNA cloning. Other antibody fragments are Fab'_2 fragments which can be produced by proteolytic digestion with pepsin.

By gene cloning it is also possible to prepare shortened antibody fragments which consist only of the variable regions of the heavy (VH) and light chain (VL). These are known as Fv fragments (fragment variable = fragment of the variable part). As covalent binding via the cystein groups of the constant chains is not possible in these Fv fragments, they are often stabilised by some other method. For this purpose the variable region of the heavy and light chains are often joined together by means of a short peptide fragment of about 10 to 30 amino acids, preferably 15 amino acids. This produces a single polypeptide chain in which VH and VL are joined together by a peptide linker. Such antibody fragments are also referred to as single chain Fv fragments (scFv). Examples of scFv antibodies are known and described, cf. for example Huston et al., 1988.

In past years various strategies have been developed for producing multimeric scFv derivatives. The intention is to produce recombinant antibodies with improved pharmacokinetic properties and increased binding avidity. In order to achieve the multimerisation of the scFv fragments they are produced as fusion proteins with multimerisation domains. The multimerisation domains may be, for example, the CH3 region of an IgG or helix structures ("coiled coil structures") such as the Leucine Zipper domains.

In other strategies the interactions between the VH and VL regions of the scFv fragment are used for multimerisation (e.g. dia-, tri- and pentabodies).

The term diabody is used in the art to denote a bivalent homodimeric scFv derivative. Shortening the peptide linker in the scFv molecule to 5 to 10 amino acids results in the formation of homodimers by superimposing VH/VL chains. The diabodies may additionally be stabilised by inserted disulphite bridges. Examples of diabodies can be found in the literature, e.g. in Perisic et al., 1994.

The term minibody is used in the art to denote a bivalent homodimeric scFv derivative. It consists of a fusion protein which contains the CH3 region of an immunoglobulin, preferably IgG, most preferably IgG1, as dimerisation region. This connects the scFv fragments by means of a hinge region, also of IgG, and a linker region. Examples of such minibodies are described by Hu et al., 1996.

The term triabody is used in the art to denote a trivalent homotrimeric scFv derivative (Kortt et al., 1997). The direct fusion of VH VL without the use of a linker sequence leads to the formation of trimers.

The fragments known in the art as mini antibodies which have a b-i, tri- or tetravalent structure are also derivatives of scFv fragments. The multimerisation is achieved by means of di-, tri- or tetrameric coiled coil structures (Pack et al., 1993 and 1995; Lovejoy et al., 1993).

Gene which codes for a fluorescent protein

In another embodiment the expression vector according to the invention contains a gene coding for a fluorescent protein functionally linked to the gene of interest. Preferably, both genes are transcribed under the control of a single heterologous promoter so that the protein/ product of interest and the fluorescent protein are coded by a bicistronic mRNA. This makes it possible

to identify cells which produce the protein/product of interest in large amounts, by means of the expression rate of the fluorescent protein.

The fluorescent protein may be, for example, a green, bluish-green, blue, yellow or other coloured fluorescent protein. One particular example is green fluorescent protein (GFP) obtained from *Aequorea victoria* or *Renilla reniformis* and mutants developed from them; cf. for example Bennet et al., 1998; Chalfie et al., 1994; WO 01/04306 and the literature cited therein.

Other fluorescent proteins and genes coding for them are described in WO 00/34318, WO 00/34326, WO 00/34526 and WO 01/27150 which are incorporated herein by reference. These fluorescent proteins are fluorophores of non-bioluminescent organisms of the species *Anthozoa*, for example *Anemonia majano, Clavularia sp., Zoanthus sp. I, Zoanthus sp. II, Discosoma striata, Discosoma sp. "red", Discosoma sp. "green", Discosoma sp. "Magenta", Anemonia sulcata*.

The fluorescent proteins used according to the invention contain in addition to the wild-type proteins natural or genetically engineered mutants and variants, fragments, derivatives or variants thereof which have for example been fused with other proteins or peptides. The mutations used may for example alter the excitation or emission spectrum, the formation of chromophores, the extinction coefficient or the stability of the protein. Moreover, the expression in mammalian cells or other species can be improved by codon optimisation. According to the invention the fluorescent protein may also be used in fusion with a selectable marker, preferably an amplifiable selectable marker such as dihydrofolate reductase (DHFR).

The fluorescence emitted by the fluorescent proteins makes it possible to detect the proteins, e.g. by throughflow cytometry with a fluorescence-activated cell sorter (FACS) or by fluorescence microscopy.

Other regulatory elements

The expression vector contains at least one heterologous promoter which allows expression of the gene of interest and preferably also of the fluorescent protein.

The term promoter denotes a polynucleotide sequence which allows and controls the transcription of the genes or sequences functionally connected therewith. A promoter contains recognition sequences for binding RNA polymerase and the initiation site for transcription (transcription initiation site). In order to express a desired sequence in a certain cell type or a host cell a suitable functional promoter must be chosen. The skilled man will be familiar with a variety of promoters from various sources, including constitutive, inducible and repressible promoters. They are deposited in databanks such as GenBank, for example, and may be obtained as separate elements or elements cloned within polynucleotide sequences from commercial or individual sources. In inducible promoters the activity of the promoter may be reduced or increased in response to a signal. One example of an inducible promoter is the tetracycline (tet) promoter. This contains tetracycline operator sequences (tetO) which can be induced by a tetracycline-regulated transactivator protein (tTA). In the presence of tetracycline the binding of tTA to tetO is inhibited. Examples of other inducible promoters are the jun, fos, metallothionein and heat shock promoter (see also Sambrook et al., 1989; Gossen et al., 1994).

Of the promoters which are particularly suitable for high expression in eukaryotes, there are for example the ubiquitin/S27a promoter of the hamster (WO 97/15664), SV 40 early promoter, adenovirus major late promoter, mouse metallothionein-I promoter, the long terminal repeat region of Rous Sarcoma Virus, the early promoter of human Cytomegalovirus. Examples of other heterologous mammalian promoters are the actin, immunoglobulin or heat shock promoter(s).

A corresponding heterologous promoter can be functionally connected to other regulatory sequences in order to increase/regulate the transcription activity in an expression cassette.

For example, the promoter may be functionally linked to enhancer sequences in order to increase the transcriptional activity. For this, one or more enhancers and/or several copies of an enhancer sequence may be used, e.g. a CMV or SV40 enhancer. Accordingly, an expression vector according to the invention, in another embodiment, contains one or more enhancers/ enhancer sequences, preferably a CMV or SV40 enhancer.

The term enhancer denotes a polynucleotide sequence which in the *cis* location acts on the activity of a promoter and thus stimulates the transcription of a gene functionally connected to this promoter. Unlike promoters the effect of enhancers is independent of position and orientation and they can therefore be positioned in front of or behind a transcription unit, within an intron or even within the coding region. The enhancer may be located both in the immediate vicinity of the transcription unit and at a considerable distance from the promoter. It is also possible to have a physical and functional overlap with the promoter. The skilled man will be aware of a number of enhancers from various sources (and deposited in databanks such as GenBank, e.g. SV40 enhancers, CMV enhancers, polyoma enhancers, adenovirus enhancers) which are available as independent elements or elements cloned within polynucleotide sequences (e.g. deposited at the ATCC or from commercial and individual sources). A number of promoter sequences also contain enhancer sequences such as the frequently used CMV promoter. The human CMV enhancer is one of the strongest enhancers identified hitherto. One example of an inducible enhancer is the metallothionein enhancer, which can be stimulated by glucocorticoids or heavy metals.

Another possible modification is, for example, the introduction of multiple Sp1 binding sites. The promoter sequences may also be combined with regulatory sequences which allow control/regulation of the transcription activity. Thus, the promoter can be made repressible/ inducible. This can be done for

example by linking to sequences which are binding sites for up- or down-regulating transcription factors. The above mentioned transcription factor Sp1, for example, has a positive effect on the transcription activity. Another example is the binding site for the activator protein AP1, which may act both positively and negatively on transcription. The activity of AP1 can be controlled by all kinds of factors such as, for example, growth factors, cytokines and serum (Faisst et al., 1992 and references therein). The transcription efficiency can also be increased by changing the promoter sequence by the mutation (substitution, insertion or deletion) of one, two, three or more bases and then determining, in a reporter gene assay, whether this has increased the promoter activity.

Basically, the additional regulatory elements include heterologous promoters, enhancers, termination and polyadenylation signals and other expression control elements. Both inducible and constitutively regulatory sequences are known for the various cell types.

"Transcription-regulatory elements" generally comprise a promoter upstream of the gene sequence to be expressed, transcription initiation and termination sites and a polyadenylation signal.

The term "transcription initiation site" refers to a nucleic acid in the construct which corresponds to the first nucleic acid which is incorporated in the primary transcript, i.e. the mRNA precursor. The transcription initiation site may overlap with the promoter sequences.

The term "transcription termination site" refers to a nucleotide sequence which is normally at the 3' end of the gene of interest or of the gene section which is to be transcribed, and which brings about the termination of transcription by RNA polymerase.

The "polyadenylation signal" is a signal sequence which causes cleavage at a specific site at the 3' end of the eukaryotic mRNA and post-transcriptional incorporation of a sequence of about 100-200 adenine nucleotides (polyA tail)

at the cleaved 3' end. The polyadenylation signal comprises the sequence AATAAA about 10-30 nucleotides upstream of the cleavage site and a sequence located downstream. Various polyadenylation elements are known such as tk polyA, SV40 late and early polyA or BGH polyA (described for example in US 5,122,458).

In a preferred embodiment of the present invention each transcription unit has a promoter or a promoter/enhancer element, a gene of interest and/or a marker gene as well as a transcription termination element. In another preferred embodiment the transcription unit contains two further translation regulatory units.

"Translation regulatory elements" comprise a translation initiation site (AUG), a stop codon and a polyA signal for each polypeptide to be expressed. For optimum expression it may be advisable to remove, add or change 5'- and/or 3'-untranslated regions of the nucleic acid sequence which is to be expressed, in order to eliminate any potentially unsuitable additional translation initiation codons or other sequences which might affect expression at the transcription or expression level. In order to promote expression, ribosomal consensus binding sites may alternatively be inserted immediately upstream of the start codon. In order to produce a secreted polypeptide the gene of interest usually contains a signal sequence which codes for a signal precursor peptide which transports the synthesised polypeptide to and through the ER membrane. The signal sequence is often but not always located at the amino terminus of the secreted protein and is cleaved by signal peptidases after the protein has been filtered through the ER membrane. The gene sequence will usually but not necessarily contain its own signal sequence. If the native signal sequence is not present a heterologous signal sequence may be introduced in known manner. Numerous signal sequences of this kind are known to the skilled man and deposited in sequence databanks such as GenBank and EMBL.

One important regulatory element according to the invention is the internal ribosomal entry site (IRES). The IRES element comprises a sequence which functionally activates the translation initiation independently of a 5'-terminal

methylguanosinium cap (CAP structure) and the upstream gene and in an animal cell allows the translation of two cistrons (open reading frames) from a single transcript. The IRES element provides an independent ribosomal entry site for the translation of the open reading frame located immediately downstream. In contrast to bacterial mRNA which may be multicistronic, i.e. it may code for numerous different polypeptides or products which are translated one after the other by the mRNA, the majority of mRNAs from animal cells are monocistronic and code for only one protein or product. In the case of a multicistronic transcript in a eukaryotic cell the translation would be initiated from the translation initiation site which was closest upstream and would be stopped by the first stop codon, after which the transcript would be released from the ribosome. Thus, only the first polypeptide or product coded by the mRNA would be produced during translation. By contrast, a multicistronic transcript with an IRES element which is functionally linked to the second or subsequent open reading frame in the transcript allows subsequent translation of the open reading frame located downstream thereof, so that two or more polypeptides or products coded by the same transcript are produced in the eukaryotic cell.

The IRES element may be of various lengths and various origins and may originate, for example, from the encephalomyocarditis virus (EMCV) or other Picorna viruses. Various IRES sequences and their use in the construction of vectors are described in the literature, cf. for example Pelletier et al., 1988; Jang et al., 1989; Davies et al., 1992; Adam et al., 1991; Morgan et al., 1992; Sugimoto et al., 1994; Ramesh et al., 1996; Mosser et al., 1997.

The gene sequence located downstream is functionally linked to the 3' end of the IRES element, i.e. the spacing is selected so that the expression of the gene is unaffected or only marginally affected or has sufficient expression for the intended purpose. The optimum permissible distance between the IRES element and the start codon of the gene located downstream thereof for sufficient expression can be determined by simple experiments by varying the spacing and determining the expression rate as a function of the spacing using reporter gene assays.

By the measures described it is possible to obtain an optimum expression cassette which is of great value for the expression of heterologous gene products. An expression cassette obtained by means of one or more such measures is therefore a further subject of the invention.

Hamster-Ubiquitin/S27a Promoter

In another embodiment the expression vector according to the invention contains the ubiquitin/S27a promoter of the hamster, preferably functionally linked to the gene of interest and even more preferably functionally linked to the gene of interest and the gene which codes for a fluorescent protein.

The ubiquitin/S27a promoter of the hamster is a powerful homologous promoter which is described in WO 97/15664. Such a promoter preferably has at least one of the following features: GC-rich sequence area, Sp1 binding site, polypyrimidine element, absence of a TATA box. Particularly preferred is a promoter which has an Sp1 binding site but no TATA box. Also preferred is a promoter which is constitutively activated and in particular is equally active under serum-containing, low-serum and serum-free cell culture conditions. In another embodiment it is an inducible promoter, particularly a promoter which is activated by the removal of serum.

A particularly advantageous embodiment is a promoter with a nucleotide sequence as contained in Fig. 5 of WO 97/15664. Particularly preferred are promoter sequences which contain the sequence from position -161 to -45 of Fig. 5.

The promoters used in the examples of the present patent specification each contain a DNA molecule with the sequence from position 1923 to 2406 of SEQ ID NO:21 of the attached sequence listing. This sequence corresponds to the fragment -372 to +111 from Fig. 5 of WO 97/15664 and represents the preferred promoter, i.e a preferred promoter should incorporate this sequence region. Another suitable promoter fragment contains the sequence from

position 2134 to 2406 (corresponding to -161 to +111 in Fig. 5 of WO 97/15664). A promoter which contains only the sequence from position 2251 to 2406 is no longer functional (corresponds to position -45 to +111 in Fig. 5 of WO 9/15664). It is possible to extend the promoter sequence in the 5' direction starting from position 2134.

It is also possible to use functional subfragments of the complete hamster ubiquitin/S27a promoter sequence as well as functional mutants/variants of the complete sequence of subfragments thereof which have been modified, for example, by substitution, insertion or deletion. Corresponding subfragments, mutants or variants are hereinafter also referred to as "modified promoters".

A modified promoter, optionally combined with other regulatory elements, preferably has a transcription activity which corresponds to that of the promoter fragment from position 1923 to 2406 of the nucleotide sequence given in SEQ ID NO:21 (-372 to +111 from Fig. 5 of WO 97/15664). A modified promoter proves to be useful for the purposes of the invention if it has a transcription activity which has at least 50%, preferably at least 80%, more preferably at least 90% and most preferably at least 100% of the activity of the 1923 to 2406 fragment (-372 to +111 fragment) in a comparative reporter gene assay. Particularly preferred are modified promoters which have a minimum sequence homology to the wild-type sequence SEQ ID NO:21 of the hamster ubiquitin/ S27a promoter of at least 80%, preferably at least 85%, preferably at least 90%, more preferably at least 95% and most preferably at least 97% and have a corresponding promoter activity in a comparative reporter gene assay.

In a corresponding comparative reporter gene assay the promoter fragments to be tested including the reference sequence are cloned in front of a promoterless reporter gene which codes, for example for luciferase, secreted alkaline phosphotase or green fluorescent protein (GFP). These constructs (promoter sequence + reporter gene) are subsequently introduced into the test cells, e.g. CHO-DG44, by transfection and the induction of the reporter

gene expression by the promoter fragment in question is determined by measuring the protein content of the reporter gene. A corresponding test is found for example in Ausubel et al., Current Protocols in Molecular Biology, 1994, updated.

The promoter sequence of the hamster ubiquitin/S27a promoter and the modified promoters, which may also include, for example, the 5' untranslated region or selected fragments thereof, and the coding region of the ubiquitin/S27a gene or selected fragments thereof, may be obtained by a skilled man with a knowledge of the sequence described in WO 97/15664 using various standard methods as described for example in Sambrook et al., 1989; Ausubel et al., 1994. Starting from the sequence described in WO 97/15664 a suitable fragment may be selected, for example, and an oligonucleotide probe containing the sequence of this fraction may be chemically synthesised. A probe of this kind may be used for example to clone the ubiquitin/S27a gene or the 5' untranslated region or other fragments thereof, for example by hybridisation from a library of the hamster genome. Using the reporter gene assay described above the skilled man is in a position to identify promoter-active fragments without any great effort and use them for the purposes of the present invention. The 5' untranslated region or special fragments thereof can easily be obtained by PCR amplification with corresponding primers from genomic DNA or a genomic library. Fragments of the 5' untranslated region may also be obtained by limited exonuclease III digestion from larger DNA fragments. Such DNA molecules may also be chemically synthesised or produced from chemically synthesised fragments by ligation.

Deletion, insertion and substitution mutants may be produced by "site-specific mutagenesis" and/or "PCR-based mutagenesis techniques". Corresponding methods are mentioned for example in Lottspeich and Zorbas 1998 Chapter 36.1 with further references.

By cross-hybridisation with probes from the 5' untranslated region of the hamster ubiquitin/S27a gene or from the S27a part of the hamster ubiquitin

S27a gene it is also possible to identify and isolate suitable promoter sequences from corresponding homologous genes of other, preferably mammalian species. Suitable techniques are described by way of example in Lottspeich and Zorbas 1998 Chapter 23. Genes are "homologous" for the purposes of the invention if their nucleotide sequence exhibits at least 70%, preferably at least 80%, preferably at least 90%, more preferably at least 95% and most preferably at least 97% conformity to the nucleotide sequence of the gene with which it is homologous.

Using the measures described above it is possible to obtain an optimised expression cassette which is highly valuable for the expression of heterologous gene products. An expression cassette obtained by one or more such measures is therefore a further object of the invention.

<u>Preparation of expression vectors according to the invention</u>

The expression vector according to the invention may theoretically be prepared by conventional methods known in the art, as described by Sambrook et al. (1989), for example. Sambrook also describes the functional components of a vector, e.g. suitable promoters (in addition to the hamster ubiquitin/S27a promoter), enhancers, termination and polyadenylation signals, antibiotic resistance genes, selectable markers, replication starting points and splicing signals. Conventional cloning vectors may be used to produce them, e.g. plasmids, bacteriophages, phagemids, cosmids or viral vectors such as baculovirus, retroviruses, adenoviruses, adeno-associated viruses and herpes simplex virus, as well as artificial chromosomes/mini chromosomes. The eukaryotic expression vectors typically also contain prokaryotic sequences such as, for example, replication origin and antibiotic resistance genes which allow replication and selection of the vector in bacteria. A number of eukaryotic expression vectors which contain multiple cloning sites for the introduction of a polynucleotide sequence are known and some may be obtained commercially from various companies such as Stratagene, La Jolla, CA, USA; Invitrogen, Carlsbad, CA, USA; Promega, Madison, WI, USA or BD Biosciences Clontech, Palo Alto, CA, USA.

The heterologous promoter, the gene of interest and the modified neomycin phosphotransferase gene and optionally the gene coding for a fluorescent protein, additional regulatory elements such as the internal ribosomal entry site (IRES), enhancers or a polyadenylation signal are introduced into the expression vector in a manner familiar to those skilled in the art. An expression vector according to the invention contains, at the minimum, a heterologous promoter, the gene of interest and a modified neomycin phosphotransferase gene. Preferably, the expression vector also contains a gene coding for a fluorescent protein. It is particularly preferred according to the invention to use a ubiquitin/S27a promoter as heterologous promoter. Particularly preferred is an expression vector in which the heterologous promoter, preferably a ubiquitin/S27a promoter, the gene of interest and the gene which codes for a fluorescent protein are functionally linked together or are functionally linked and the neomycin phosphotransferase gene is located in the same or in a separate transcription unit.

Within the scope of the present description the term "functional linking" or "functionally linked" refers to two or more nucleic acid sequences or partial sequences which are positioned so that they can perform their intended function. For example, a promoter/enhancer is functionally linked to a coding gene sequence if it is able to control or modulate the transcription of the linked gene sequence in the cis position. Generally, but not necessarily, functionally linked DNA sequences are close together and, if two coding gene sequences are linked or in the case of a secretion signal sequence, in the same reading frame. Although a functionally linked promoter is generally located upstream of the coding gene sequence it does not necessarily have to be close to it. Enhancers need not be close by either, provided that they assist the transcription of the gene sequence. For this purpose they may be both upstream and downstream of the gene sequence, possibly at some distance from it. A polyadenylation site is functionally linked to a gene sequence if it is positioned at the 3' end of the gene sequence in such a way that the transcription progresses via the coding sequence to the polyadenylation signal. Linking may take place according to conventional recombinant

methods, e.g. by the PCR technique, by ligation at suitable restriction cutting sites or by splicing. If no suitable restriction cutting sites are available synthetic oligonucleotide linkers or adaptors may be used in a manner known *per se*. According to the invention the functional linking preferably does not take place via intron sequences.

In one of the embodiments described, the heterologous promoter, preferably a ubiquitin/S27a promoter, the gene of interest and the gene coding for a fluorescent protein are functionally linked together. This means for example that both the gene of interest and the gene coding for a fluorescent protein are expressed starting from the same heterologous promoter.

In a particularly preferred embodiment the functional linking takes place via an IRES element, so that a bicistronic mRNA is synthesised from both genes. The expression vector according to the invention may additionally contain enhancer elements which act functionally on one or more promoters. Particularly preferred is an expression vector in which the heterologous promoter, preferably the ubiquitin/S27a promoter or a modified form thereof, is linked to an enhancer element, e.g. an SV40 enhancer or a CMV enhancer element.

Fundamentally, the expression of the genes within an expression vector may take place starting from one or more transcription units. The term transcription unit is defined as a region which contains one or more genes to be transcribed. The genes within a transcription unit are functionally linked to one another in such a way that all the genes within such a unit are under the transcriptional control of the same promoter or promoter/ enhancer. As a result of this transcriptional linking of genes, more than one protein or product can be transcribed from a transcription unit and thus expressed. Each transcription unit contains the regulatory elements which are necessary for the transcription and translation of the gene sequences contained therein. Each transcription unit may contain the same or different regulatory elements. IRES elements or introns may be used for the functional linking of the genes within a transcription unit.

The expression vector may contain a single transcription unit for expressing the gene of interest, the modified NPT gene and optionally the gene which codes for the fluorescent protein. Alternatively, these genes may also be arranged in two or more transcription units. Various combinations of the genes within a transcription unit are possible. In another embodiment of the present invention more than one expression vector consisting of one, two or more transcription units may be inserted in a host cell by cotransfection or in successive transfections in any desired order. Any combination of regulatory elements and genes on each vector can be selected provided that adequate expression of the transcription units is ensured. If necessary, other regulatory elements and genes, e.g. additional genes of interest or selectable markers, may be positioned on the expression vectors.

Accordingly, an expression vector according to the invention containing a gene of interest and a gene which codes for a modified neomycin phosphotransferase may contain both genes in one or in two separate transcription units. Each transcription unit can transcribe and express one or more gene products. If both genes are contained in one transcription unit they are under the control of the same promoter or promoter/ enhancer, while preferably an IRES element is used to ensure the functional linking of all the components. If the gene which codes for modified neomycin phosphotransferase and the gene of interest are contained in two separate transcription units, they may be under the control of the same or different promoters/enhancers. However, preferably, a weaker heterologous promoter, e.g. SV40 early promoter, is used for the modified NPT gene and preferably no enhancer is used. Expression vectors with two separate transcription units are preferred within the scope of the invention. One (bicistronic) transcription unit contains the gene of interest and optionally a gene coding for a fluorescent protein, while the other transcription unit contains the modified NPT gene. Preferably, each transcription unit is limited at the 3' end by a sequence which codes for a polyA signal, preferably BGH polyA or SV40 polyA.

Also preferred according to the invention are those expression vectors which instead of the gene of interest have only a multiple cloning site which allows the cloning of the gene of interest via recognition sequences for restriction endonucleases. Numerous recognition sequences for all kinds of restriction endonucleases as well as the associated restriction endonucleases are known from the prior art. Preferably, sequences are used which consist of at least six nucleotides as recognition sequence. A list of suitable recognition sequences can be found for example in Sambrook et al., 1989.

Host Cells

For transfection with the expression vector according to the invention eukaryotic host cells are used, preferably mammalian cells and more particularly rodent cells such as mouse, rat and hamster cell lines. The successful transfection of the corresponding cells with an expression vector according to the invention results in transformed, genetically modified, recombinant or transgenic cells, which are also the subject of the present invention.

Preferred host cells for the purposes of the invention are hamster cells such as BHK21, BHK TK^{-}, CHO, CHO-K1, CHO-DUKX, CHO-DUKX B1 and CHO-DG44 cells or derivatives/descendants of these cell lines. Particularly preferred are CHO-DG44, CHO-DUKX, CHO-K1 and BHK21 cells, particularly CHO-DG44 and CHO-DUKX cells. Also suitable are myeloma cells from the mouse, preferably NS0 and Sp2/0 cells and derivatives/descendants of these cell lines.

Examples of hamster and mouse cells which can be used according to the invention are given in Table 2 that follows. However, derivatives and descendants of these cells, other mammalian cells including but not restricted to cell lines of humans, mice, rats, monkeys, rodents, or eukaryotic cells, including but not restricted to yeast, insect and plant cells, may also be used as host cells for the production of biopharmaceutical proteins.

Table 2: Hamster and Mouse Production Cell Lines

Cell line	Accession Number
NS0	ECASS No. 85110503
Sp2/0-Ag14	ATCC CRL-1581
BHK21	ATCC CCL-10
BHK TK$^-$	ECACC No. 85011423
HaK	ATCC CCL-15
2254-62.2 (BHK-21-derivative)	ATCC CRL-8544
CHO	ECACC No. 8505302
CHO-K1	ATCC CCL-61
CHO-DUKX (= CHO duk$^-$, CHO/dhfr$^-$)	ATCC CRL-9096
CHO-DUKX B1	ATCC CRL-9010
CHO-DG44	Urlaub et al; Cell 32[2], 405-412, 1983
CHO Pro-5	ATCC CRL-1781
V79	ATCC CCC-93
B14AF28-G3	ATCC CCL-14
CHL	ECACC No. 87111906

The transfection of the eukaryotic host cells with a polynucleotide or one of the expression vectors according to the invention is carried out by conventional methods (Sambrook et al., 1989; Ausubel et al., 1994). Suitable methods of transfection include for example lyposome-mediated transfection, calcium phosphate co precipitation, electroporation, polycation- (e.g. DEAE dextran)-mediated transfection, protoplast fusion, microinjection and viral infections. According to the invention stable transfection is preferably carried out in which the constructs are either integrated into the genome of the host cell or an artificial chromosome/minichromosome, or are episomally contained in stable manner in the host cell. The transfection method which gives the

optimum transfection frequency and expression of the heterologous gene in the host cell in question is preferred. By definition, every sequence or every gene inserted in a host cell is referred to as a "heterologous sequence" or "heterologous gene" in relation to the host cell. This applies even if the sequence to be introduced or the gene to be introduced is identical to an endogenous sequence or an endogenous gene of the host cell. For example, a hamster actin gene introduced into a hamster host cell is by definition a heterologous gene.

According to the invention, recombinant mammalian cells, preferably rodent cells, most preferably hamster cells such as CHO or BHK cells which have been transfected with one of the expression vectors according to the invention described herein are preferred.

In the recombinant production of heteromeric proteins such as e.g. monoclonal antibodies (mAb), the transfection of suitable host cells can theoretical be carried out by two different methods. mAb's of this kind are composed of a number of subunits, the heavy and light chains. Genes coding for these subunits may be accommodated in independent or multicistronic transcription units on a single plasmid with which the host cell is then transfected. This is intended to secure the stoichiometric representation of the genes after integration into the genome of the host cell. However, in the case of independent transcriptional units it must hereby be ensured that the mRNAs which encode the different proteins display the same stability and transcriptional and translational efficiency. In the second case, the expression of the genes take place within a multicistronic transcription unit by means of a single promoter and only one transcript is formed. By using IRES elements, a highly efficient internal translation initiation of the genes is obtained in the second and subsequent cistrons. However, the expression rates for these cistrons are lower than that of the first cistron, the translation initiation of which, by means of a so-called "cap"-dependent pre-initiation complex, is substantially more efficient than IRES-dependent translation initiation. In order to achieve a truly equimolar expression of the cistrons, additional inter-

cistronic elements may be introduced, for example, which ensure uniform expression rates in conjunction with the IRES elements (WO 94/05785).

Another possible way of simultaneously producing a number of heterologous proteins, which is preferred according to the invention, is cotransfection, in which the genes are separately integrated in different expression vectors. This has the advantage that certain proportions of genes and gene products with one another can be adjusted, thereby balancing out any differences in the mRNA stability and in the efficiency of transcription and translation. In addition, the expression vectors are more stable because of their small size and are easier to handle both during cloning and during transfection.

In one particular embodiment of the invention, therefore, the host cells are additionally transfected, preferably cotransfected, with one or more vectors having genes which code for one or more other proteins of interest. The other vector or vectors used for the cotransfection code, for example, for the other protein or proteins of interest under the control of the same promoter/enhancer combination and for at least one other selectable marker, e.g. dihydrofolate reductase.

According to the invention the host cells are preferably established, adapted and cultivated under serum-free conditions, optionally in media which are free from animal proteins/peptides. Examples of commercially obtainable media include Ham's F12 (Sigma, Deisenhofen, DE), RPMI-1640 (Sigma), Dulbecco's Modified Eagle's Medium (DMEM; Sigma), Minimal Essential Medium (MEM; Sigma), Iscove's Modified Dulbecco's Medium (IMDM; Sigma), CD-CHO (Invitrogen, Carlsbad, CA, USA), CHO-S-SFMII (Invitrogen), serum-free CHO-Medium (Sigma) and protein-free CHO-Medium (Sigma). Each of these media may optionally be supplemented with various compounds, e.g. hormones and/or other growth factors (e.g. insulin, transferrin, epidermal growth factor, insulin-like growth factor), salts (e.g. sodium chloride, calcium, magnesium, phosphate), buffers (e.g. HEPES), nucleosides (e.g. adenosine, thymidine), glutamine, glucose or other equivalent nutrients, antibiotics and/or trace elements. Although serum-free

media are preferred according to the invention, the host cells may also be cultivated using media which have been mixed with a suitable amount of serum. In order to select genetically modified cells which express one or more selectable marker genes, one or more selecting agents are added to the medium.

The term "selecting agent" refers to a substance which affects the growth or survival of host cells with a deficiency for the selectable marker gene in question. Within the scope of the present invention, geneticin (G418) is preferably used as the medium additive for the selection of heterologous host cells which carry a modified neomycin phosphotransferase gene. Preferably, G418 concentrations of between 100 and 800 μg/ml of medium are used, most preferably 300 to 400 μg/ml of medium. If the host cells are to be transfected with a number of expression vectors, e.g. if several genes of interest are to be separately introduced into the host cell, they generally have different selectable marker genes.

A selectable marker gene is a gene which allows the specific selection of cells which contain this gene by the addition of a corresponding selecting agent to the cultivation medium. As an illustration, an antibiotic resistance gene may be used as a positive selectable marker. Only cells which have been transformed with this gene are able to grow in the presence of the corresponding antibiotic and are thus selected. Untransformed cells, on the other hand, are unable to grow or survive under these selection conditions. There are positive, negative and bifunctional selectable markers. Positive selectable markers permit the selection and hence enrichment of transformed cells by conferring resistance to the selecting agent or by compensating for a metabolic or catabolic defect in the host cell. By contrast, cells which have received the gene for the selectable marker can be selectively eliminated by negative selectable markers. An example of this is the thymidine kinase gene of the Herpes Simplex virus, the expression of which in cells with the simultaneous addition of acyclovir or gancyclovir leads to the elimination thereof. The selectable markers used in this invention, including the

amplifiable selectable markers, include genetically modified mutants and variants, fragments, functional equivalence, derivatives, homologues and fusions with other proteins or peptides, provided that the selectable marker retains its selective qualities. Such derivatives display considerable homology in the amino acid sequence in the regions or domains which are deemed to be selective. The literature describes a large number of selectable marker genes including bifunctional (positive/negative) markers (see for example WO 92/08796 and WO 94/28143). Examples of selectable markers which are usually used in eukaryotic cells include the genes for aminoglycoside phosphotransferase (APH), hygromycine phosphostransferase (HYG), dihydrofolate reductase (DHFR), thymidine kinase (TK), glutamine synthetase, asparagin synthetase and genes which confer resistance to neomycin (G418), puromycin, histidinol D, belomycin, phleomycin and zeocin.

It is also possible to select transformed cells by fluorescence-activated cell sorting (FACS). For this, bacterial β-galactosidase, cell surface markers or fluorescent proteins may be used (e.g. green fluorescent protein (GFP) and the variants thereof from *Aequorea victoria* and *Renilla reniformis* or other species; red fluorescent protein and proteins which fluoresce in other colours and their variants from non-bioluminescent organisms such as e.g. *Discosoma sp., Anemonia sp., Clavularia sp., Zoanthus sp.*) for the selection of transformed cells.

Gene expression and selection of high-producing host cells

The term gene expression relates to the transcription and/or translation of a heterologous gene sequence in a host cell. The expression rate can be generally determined, either on the basis of the quantity of corresponding mRNA which is present in the host cell or on the basis of the quantity of gene product produced which is encoded by the gene of interest. The quantity of mRNA produced by transcription of a selected nucleotide sequence can be determined for example by northern blot hybridisation, ribonuclease-RNA-protection, *in situ* hybrisation of cellular RNA or by PCR methods (Sambrook et al., 1989; Ausubel et al., 1994). Proteins which are encoded by a selected

nucleotide sequence can also be determined by various methods such as, for example, ELISA, western blot, radioimmunoassay, immunoprecipitation, detection of the biological activity of the protein or by immune staining of the protein followed by FACS analysis (Sambrook et al., 1989; Ausubel et al., 1994).

The terms "high expression level (or rate), high expression, increased expression or high productivity" refer to the long-lasting and sufficiently high expression or synthesis of a heterologous sequence introduced into a host cell, e.g. of a gene coding for a therapeutic protein. Increased or high expression or a high expression level or rate or a high productivity are present if a cell according to the invention is cultivated by one of the methods according to the invention described here and if this cell produces at least more than roughly 5 pg of the desired gene product per day (5 pg/day/cell). Increased or high expression or a high expression or rate or a high productivity are also present if the cell according to the invention produces at least more than roughly 10 pg of the desired gene produce per day (10 pg/day/cell). Increased or high expression or a high expression level or rate or high productivity are present in particular if the cell according to the invention produces at least more than roughly 15 pg of the desired gene product per day (15 pg/ day/cell). Increased or high expression or a high expression level or rate or high productivity are present in particular if the cell according to the invention produces at least more than roughly 20 pg of the desired gene product per day (20 pg/day/cell). Particularly increased or high expression or a particularly high expression level or rate or particularly high productivity are present if the cell according to the invention produces at least more than roughly 30 pg of the desired gene product per day (30 pg/day/cell).

High or increased expression, high productivity or a high expression level or rate can be achieved both by using one of the expression vectors according to the invention and by the use of one of the processes according to the invention.

For example, by coexpression of the gene of interest and a modified NPT gene it is possible to select and identify cells which express the heterologous gene to a high degree. Compared with wtNPT, modified NPT allows more efficient selection of stably transfected host cells with high expression of the heterologous gene of interest.

The present invention thus also relates to a process for expressing at least one gene of interest in recombinant mammalian cells, characterised in that (i) a pool of mammalian cells is transfected with at least one gene of interest and one gene for a modified neomycin phosphotransferase which compared with the wild-type neomycin phosphotransferase has only 10 to 50% of its activity, preferably 20 to 50%, more particularly 25% to 37% and more preferably 25 to 32%; (ii) the cells are cultivated under conditions which allow expression of the gene or genes of interest and the modified neomycin phosphotransferase gene; (iii) the mammalian cells are cultivated in the presence of at least one selecting agent, preferably G418, which acts selectively on the growth of the mammalian cells, and gives preference to the growth of those cells which express the modified neomycin phosphotransferase gene; and (iv) the protein or proteins of interest is or are obtained from the mammalian cells or the culture supernatant. Preferably recombinant mammalian cells are used which have been transfected with an expression vector according to the invention.

The invention also relates to a process for selecting recombinant mammalian cells which express at least one gene of interest, wherein (i) a pool of mammalian cells is transfected with at least one gene of interest and a gene for a modified neomycin phosphotransferase which by comparison with wild-type neomycin phosphotransferase has only 10 to 50% of the activity, preferably 20 to 50%, particularly 25 to 37% and more preferably 25 to 32%; (ii) the mammalian cells are cultivated under conditions which allow expression of the gene or genes of interest and the modified neomycin phosphotransferase gene; and (iii) the mammalian cells are cultivated in the presence of at least one selecting agent, preferably G418, which acts selectively on the growth of the mammalian cells and gives preference to the

growth of those cells which express the modified neomycin phosphotransferase gene.

Particularly preferred are processes for expressing at least one gene of interest and for selecting recombinant cells which express a corresponding gene of interest if a modified NPT gene described in more detail in this application is used, particularly if a modified NPT gene is used which by comparison with the wild-type gene codes for an aspartic acid at amino acid position 182 or for glycine at amino acid position 227. It is particularly preferred to use the Asp227Gly mutant.

The selection of the cells which express a gene of interest and a modified NPT gene is carried out for example by adding G418 as selecting agent. However, it is also possible to use other aminoglycoside antibiotics such as neomycin or kanamycin. The cells according to the invention are preferably cultivated and selected in 200 to 800 μg of G418 per mL of culture medium. It has proved particularly preferable to add 300 to 700 μg of G418 per mL of culture medium. The addition of roughly 400 μg of G418 per mL of culture medium is the most preferred embodiment. Using such a method it is possible to select recombinant cells with a particularly high expression rate. By comparison with the use of wtNPT as selectable marker the cells exhibited a productivity increased by a factor of 2.4 in the case of the Glu182Asp mutant and productivity increased by a factor of up to 4.1 in the case of the Asp227Gly mutant after selection with 400 μg of G418 per mL of culture medium.

The corresponding processes may be combined with a FACS-assisted selection of recombinant host cells which contain, as additional selectable marker, one or more fluorescent proteins (e.g. GFP) or a cell surface marker. Other methods of obtaining increased expression, and a combination of different methods may also be used, are based for example on the use of (artificial) transcription factors, treatment of the cells with natural or synthetic agents for up-regulating endogenous or heterologous gene expression,

improving the stability (half-life) of mRNA or the protein, improving the initiation of mRNA translation, increasing the gene dose by the use of episomal plasmids (based on the use of viral sequences as replication origins, e.g. SV40, polyoma, adenovirus, EBV or BPV), the use of amplification-promoting sequences (Hemann et al., 1994) or in vitro amplification systems based on DNA concatemers (Monaco et al., 1996).

Coupled transcription of the gene of interest and the gene which codes for the fluorescent protein has proved particularly effective in conjunction with the use of a modified NPT gene as selectable marker. The resulting bicistronic mRNA expresses both the protein/product of interest and the fluorescent protein. On the basis of this coupling of the expression of the protein of interest and the fluorescent protein it is easily possible according to the invention to identify and isolate high-producing recombinant host cells by means of the fluorescent protein expressed, e.g. by sorting using fluorescence activated cell sorting equipment (FACS).

The selection of recombinant host cells which exhibit high vitality and an increased expression rate of the desired gene product is a multistage process. The host cells which have been transfected with the expression vector according to the invention or optionally cotransfected with another vector, for example, are cultivated under conditions which permit the selection of cells expressing the modified NTP, e.g. by cultivation in the presence of a selecting agent such as G418 in concentrations of 100, 200, 400, 600, 800 µg or more of G418/mL of culture medium. Then the corresponding cells are investigated at least for the expression of the gene which codes for a fluorescent protein and is coupled to the gene of interest, in order to identify and sort out the cells/cell population which exhibit the highest expression rates of fluorescent protein. Preferably, only the cells which belong to the 10% of cells with the highest expression rate of fluorescent protein are sorted out and further cultivated. In practice this means that the brightest 10% of the fluorescent cells are sorted out and further cultivated. Accordingly, the brightest 5%, preferably the brightest 3% or even the brightest 1% of the fluorescent cells of a cell mixture can also be sorted out and replicated. In a

particularly preferred embodiment only the brightest 0.5% or the brightest 0.1% of the fluorescent cells are sorted out and replicated.

The selection step may be carried out on cell pools or using pre-sorted cell pools/cell clones. One or more, preferably two or more and especially three or more sorting steps may be carried out, while between the individual sorting steps the cells may be cultivated and replicated for a specific length of time, e.g. roughly two weeks in the case of pools.

The present invention thus relates to a process for obtaining and selecting recombinant mammalian cells which express at least one heterologous gene of interest, characterised in that (i) recombinant mammalian cells are transfected with an expression vector according to the invention; (ii) the transfected cells are cultivated under conditions which allow expression of the gene or genes of interest, the gene coding for a fluorescent protein and the modified neomycin phosphotransferase gene; (iii) the mammalian cells are cultivated in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells and gives preference to the growth of those cells which express the modified neomycin phosphotransferase gene; and (iv) the mammalian cells which exhibit a particularly high expression of the fluorescent gene are sorted out by flow-cytometric analysis. If desired steps (ii) to (iv) may be repeated once or several times with the cells obtained in step (iv).

A corresponding process is preferred which is characterised in that the sorted mammalian cells have an average specific productivity of more than 5 pg of the desired gene product or products per day and per cell (5 pg/day/cell), preferably greater than 10 pg/day/cell, more preferably greater than 20 pg/day/cell.

Also preferred according to the invention is a process in which suitably sorted cells are replicated and used to prepare the encoded gene product of interest. For this, the selected high producing cells are preferably cultivated in a serum-free culture medium and preferably in suspension culture under

conditions which allow expression of the gene of interest. The protein/product of interest is preferably obtained from the cell culture medium as a secreted gene product. If the protein is expressed without a secretion signal, however, the gene product may also be isolated from cell lysates. In order to obtain a pure homogeneous product which is substantially free from other recombinant proteins and host cell proteins, conventional purification procedures are carried out. First of all, cells and cell debris are removed from the culture medium or lysate. The desired gene product can then be freed from contaminating soluble proteins, polypeptides and nucleic acids, e.g. by fractionation on immunoaffinity and ion exchange columns, ethanol precipitation, reversed phase HPLC or chromatography on Sephadex, silica or cation exchange resins such as DEAE. Methods which result in the purification of a heterologous protein expressed by recombinant host cells are known to the skilled man and described in the literature, e.g. by Harris et al., 1995 and Scopes 1988.

Amplifiable Selectable Marker Gene

In addition, the cells according to the invention may optionally also be subjected to one or more gene amplification steps in which they are cultivated in the presence of a selecting agent which leads to amplification of an amplifiable selectable marker gene. This step may be carried out both with cells which have not yet been sorted and also with cells which have already been pre-sorted once or several times.

The prerequisite is that the host cells are additionally transfected with a gene which codes for an amplifiable selectable marker. It is conceivable for the gene which codes for an amplifiable selectable marker to be present on one of the expression vectors according to the invention or to be introduced into the host cell by means of another vector.

The amplifiable selectable marker gene usually codes for an enzyme which is needed for the growth of eukaryotic cells under certain cultivation conditions. For example, the amplifiable selectable marker gene may code for

dihydrofolate reductase (DHFR). In this case the gene is amplified if a host cell transfected therewith is cultivated in the presence of the selecting agent methotrexate (MTX).

The following Table 1 gives examples of other amplifiable selectable marker genes and the associated selecting agents which may be used according to the invention, which are described in an overview by Kaufman, Methods in Enzymology, 185:537-566 (1990).

Table 1: Amplifiable selectable marker genes

Amplifiable selectable marker gene	Accession number	Selecting agent
dihydrofolate reductase	M19869 (hamster) E00236 (mouse)	methotrexate (MTX)
metallothionein	D10551 (hamster) M13003 (human) M11794 (rat)	cadmium
CAD (carbamoylphosphate synthetase : aspartate transcarbamylase: dihydroorotase)	M23652 (hamster) D78586 (human)	N-phosphoacetyl-L-aspartate
adenosine-deaminase	K02567 (human) M10319 (mouse)	Xyl-A- or adenosine, 2´deoxycoformycin
AMP (adenylate)-deaminase	D12775 (human) J02811 (rat)	adenine, azaserin, coformycin
UMP-synthase	J03626 (human)	6-azauridine, pyrazofuran
IMP 5´-dehydrogenase	J04209 (hamster) J04208 (human) M33934 (mouse)	mycophenolic acid
xanthine-guanin-phosphoribosyltransferase	X00221 (E. coli)	mycophenolic acid with limiting xanthine
mutant HGPRTase or mutant thymidine-kinase	J00060 (hamster) M13542, K02581 (human) J00423, M68489(mouse) M63983 (rat) M36160 (Herpes virus)	hypoxanthine, aminopterine and thymidine (HAT)
thymidylate-synthetase	D00596 (human) M13019 (mouse) L12138 (rat)	5-fluorodeoxyuridine
P-glycoprotein 170 (MDR1)	AF016535 (human) J03398 (mouse)	several drugs, e.g. adriamycin, vincristin, colchicine

ribonucleotide reductase	M124223, K02927 (mouse)	aphidicoline
glutamine-synthetase	AF150961 (hamster) U09114, M60803 (mouse) M29579 (rat)	methionine sulphoximine (MSX)
asparagine-synthetase	M27838 (hamster) M27396 (human) U38940 (mouse) U07202 (rat)	β-aspartylhydroxamate, albizziin, 5´azacytidine
argininosuccinate-synthetase	X01630 (human) M31690 (mouse) M26198 (bovine)	canavanin
ornithine-decarboxylase	M34158 (human) J03733 (mouse) M16982 (rat)	α-difluoromethylornithine
HMG-CoA-reductase	L00183,M12705 (hamster) M11058 (human)	compactin
N-acetylglucosaminyl-transferase	M55621 (human)	tunicamycin
threonyl-tRNA-synthetase	M63180 (human)	borrelidin
Na^+K^+-ATPase	J05096 (human) M14511 (rat)	ouabain

According to the invention the amplifiable selectable marker gene used is preferably a gene which codes for a polypeptide with the function of DHFR, e.g. for DHFR or a fusion protein from the fluorescent protein and DHFR. DHFR is necessary for the biosynthesis of purines. Cells which lack the DHFR genes cannot grow in purine-deficient medium. The DHFR gene is therefore a useful selectable marker for selecting and amplifying genes in cells cultivated in purine-free medium. The selecting medium used in conjunction with the DHFR gene is methotrexate (MTX).

The present invention therefore includes a method of preparing and selecting recombinant mammalian cells which contains the following steps: (i) transfection of the host cells with genes which code at least for a protein/ product of interest, a modified neomycin phosphotransferase and DHFR; (ii) cultivation of the cells under conditions which allow expression of the various genes; and (iii) the amplification of the co-integrated genes by cultivating the cells in the presence of a selecting agent which allows the amplification of at

least the amplifiable selectable marker gene such as methotrexate. Preferably the transfected cells are cultivated in hypoxanthine/thymidine-free medium in the absence of serum and with the addition of increasing concentrations of MTX. Preferably the concentration of MTX in the first amplification step is at least 100 nM. The concentration of MTX may, however, also be at least 250 nM and be increased step by step to 1 μM. In individual cases concentrations of more than 1 μM may be used, e.g. 2 μM.

In the present case host cells which had been identified and sorted using a fluorescence activated cell sorting device (FACS) were cultivated in a gene amplification step in the presence of 100 nM MTX. In this way it was possible to increase productivities substantially to more than 20 pg of gene product per cell and per day.

If desired the host cells may be subjected to one or more gene amplification steps in order to increase the copy number of at least the gene of interest and the amplifiable selectable marker gene. According to the invention the high productivity which can be achieved is not tied to a high number of gene copies but to effective pre-selection by means of neomycin phosphotransferase-mediated resistance to aminoglycoside antibiotics such as neomycin, kanamycin and G418. It is therefore possible to reduce the number of gene amplification steps required and to carry out only a single gene amplification, for example.

In a further embodiment the present invention thus also relates to processes for obtaining and selecting recombinant mammalian cells which express at least one heterologous gene of interest and are characterised in that (i) recombinant mammalian cells are transfected with an expression vector according to the invention and the gene for an amplifiable selectable marker gene; (ii) the mammalian cells are cultivated under conditions which allow expression of the gene or genes of interest, the modified neomycin phosphotransferase gene and the gene which codes for a fluorescent protein; (iii) the mammalian cells are cultivated in the presence of at least one

selecting agent which acts selectively on the growth of mammalian cells and gives preference to the growth of those cells which express the neomycin phosphotransferase gene; (iv) the mammalian cells which exhibit high expression of the fluorescent protein are sorted out by flow-cytometric analysis; (v) the sorted cells are cultivated under conditions under which the amplifiable selectable marker gene is expressed; and (vi) a selecting agent is added to the culture medium which results in the amplification of the amplifiable selectable marker gene.

Particularly preferred is a corresponding process in which the modified neomycin phosphotransferase genes described in this invention are used. Also preferred is a process in which only one amplification step is carried out. Also preferred is a corresponding process which leads to recombinant mammalian cells which exhibit an average specific productivity of more than 20 pg of the desired gene product or products per cell and per day.

Mammalian cells, preferably mouse myeloma and hamster cells, are preferred host cells for the use of DHFR as an amplifiable selectable marker. The cell lines CHO-DUKX (ATCC CRL-9096) and CHO-GD44 (Urlaub et al., 1983) are particularly preferred as they have no DHFR activity of their own, as a result of mutation. In order to be able to use the DHFR-induced amplification in other cell types as well which have their own endogenous DHFR activity, it is possible to use in the transfection process a mutated DHFR gene which codes for a protein with reduced sensitivity to methotrexate (Simonson et al., 1983; Wigler et al., 1980; Haber et al., 1982).

The DHFR marker is particularly suitable for the selection and subsequent amplification when using DHFR negative basic cells such as CHO-DG44 or CHO-DUKX, as these cells do not express endogenous DHFR and therefore do not grow in purine-free medium. Consequently, the DHFR gene may be used here as a dominant selectable marker and the transformed cells are selected in hypoxanthine/ thymidine-free medium.

The invention is described more fully hereinafter with reference to some non-restrictive examples.

EXAMPLES

Abbreviations

AP:	alkaline phosphatase
Asp (=D):	aspartic acid
bp:	base pair
BSA:	bovine serum albumin
CHO:	Chinese Hamster Ovary
dhfr:	dihydrofolate-reductase
DMSO:	dimethylsulphoxide
ELISA:	enzyme-linked immunosorbent assay
FACS:	fluorescence-activated cell sorter
FITC:	fluoresceine-isothiocyanate
GFP:	green fluorescent protein
Glu (=E):	glutamic acid
Gly (=G):	glycine
HBSS:	Hanks Balanced Salt Solution
HT:	hypoxanthine/thymidine
IRES:	internal ribosomal entry site
kb:	kilobase
mAb:	monoclonal antibody
MCP-1:	monocyte chemoattractant protein-1
MTX:	methotrexate
MW:	mean value
NPT:	neomycin-phosphotransferase
PCR:	polymerase chain reaction
PBS:	phosphate buffered saline
WT:	wild-type

Methods

1. Cell culture and Transfection

The cells CHO-DG44/dhfr$^{-/-}$ (Urlaub et al., 1983) were permanently cultivated as suspension cells in serum-free CHO-S-SFMII medium supplemented with hypoxanthine and thymidine (Invitrogen GmbH, Karlsruhe, DE) in cell culture flasks at 37°C in a damp atmosphere and 5% CO_2. The cell counts and viability were determined with a CASY1 Cell Counter (Schaerfe System, DE) or by tryptan blue staining and the cells were then seeded in a concentration of 1 – 3 x10^5/mL and run every 2 – 3 days.
Lipofectamine Plus Reagent (Invitrogen GmbH) was used for the transfection of CHO-DG44. For each transfection mixture a total of 1 µg of plasmid-DNA, 4 µL of lipofectamine and 6 µL of Plus reagent were mixed together according to the manufacturer's instructions and added in a volume of 200 µL to 6 x10^5 exponentially growing CHO-DG44 cells in 0.8 mL of HT-supplemented CHO-S-SFMII medium. After three hours' incubation at 37°C in a cell incubator 2 mL of HT-supplemented CHO-S-SFMII medium was added. For the NPT-based selection the cells were transferred 2 days after transfection into HT-supplemented CHO-S-SFMII with G418 (Invitrogen), changing the medium every 3 to 4 days. As a rule, 400 µg/mL of G418 were added for the selection, while in some experimental series the concentration was also lowered to 200 µg/mL or raised to 500, 600 or 800 µg/mL. In a DHFR- and NPT-based selection in the case of co-transfection in which one expression vector contained a DHFR and the other expression vector contained a neomycin-phosphotransferase selectable marker, the cells were transferred 2 days after transfection into CHO-S-SFMII medium without the addition of hypoxanthine and thymidine and G418 (Invitrogen) was also added to the medium in a concentration of 400µg/mL.
A DHFR-based gene amplification of the integrated heterologous genes was obtained by the addition of the selecting agent MTX (Sigma, Deisenhofen, DE) in a concentration of 5 – 2000 nM to the HT-free CHO-S-SFMII medium.

2. Expression vectors

To analyse the expression, eukaryotic expression vectors were used which are based on the pAD-CMV vector (Werner et al., 1998) and mediate the constitutive expression of a heterologous gene by the combination of CMV enhancer/hamster ubiquitin/S27a promoter (WO 97/15664). While the base vector pBID contains the dhfr-minigene which acts as an amplifiable selectable marker (cf e.g. EP-0-393-438), in the vector pBIN the dhfr-minigene has been replaced by a neomycin-phosphotransferase resistance gene (Fig.1). For this purpose the selectable marker neomycin-phosphotransferase, including SV40 early promoter and TK-polyadenylation signal, was isolated from the commercial plasmid pBK-CMV (Stratagene, La Jolla, CA, USA) as a 1640 bp Bsu36I fragment. After a reaction to fill in the ends of the fragment with Klenow-DNA-polymerase the fragment was ligated with the 3750 bp Bsu36I/StuI fragment of the vector pBID, which was also treated with Klenow-DNA-polymerase.

In the bicistronic base vector pBIDG (Fig. 1) the IRES-GFP gene region was isolated from the vector pIRES2-EGFP (Clontech, Palo Alto, CA, USA) and brought under the control of the CMV enhancer/promoter in the vector pBID so that the multiple cloning site between the promoter region and IRES-element was retained. The following procedure was used. In a PCR mutagenesis in which the plasmid pIRES2-EGFP acted as the template, on the one hand the HindIII cutting site AAGCTT within the IRES sequence was converted into the sequence ATGCTT by the use of mutagenic primers and thus eliminated. On the other hand an XbaI cutting site was inserted by means of a primer with complementarity to the 5´end of the IRES sequence or a SpeI cutting site was introduced by means of a primer with complementarity to the 3´end of the GFP sequence. The resulting PCR fragment, which contained the complete IRES and GFP sequence, was digested with XbaI and SpeI and cloned into the singular XbaI cutting site at the 3´end of the multiple cloning site of the vector pBID. In the same way the IRES-GFP gene region from the vector pIRES2-EGFP was brought under the control of the CMV enhancer/hamster ubiquitin/S27a promoter in the vector pKS-N8 (preparation

described in Example 1). This produced the bicistronic base vector pBING (Fig.1).

The human MCP-1 cDNA (Yoshimura et al., 1989) was cloned as a 0. 3 kb HindIII/EcoRI fragment into the corresponding cutting sites of the vector pBIN, resulting in the vector pBIN-MCP1 (Fig. 2).

In order to express a monoclonal humanised IgG2 antibody the heavy chain was cloned as a 1.5 kb BamHI/HindIII fragment into the vector pBID or pBIDG digested with BamHI and HindIII, to obtain the vector pBID-HC or pBIDG-HC (Fig. 2). The light chain on the other hand was cloned as a 0.7 kb BamHI/HindIII fragment into the corresponding cutting sites of the vector pBIN or pBING, producing the vector pBIN-LC (Fig. 2) or pBING-LC (Fig. 4).

3. FACS

The flow-cytometric analyses and sorting were carried out with a Coulter Epics Altra device. The FACS is fitted with a helium-argon laser with an excitation wavelength of 488 nm. The fluorescence intensity is absorbed at a wavelength suited to the fluorescence protein and process by means of the attached software Coulter Expo32. The sorting was carried out at a rate of 8000 – 10000 events/second. The suspended cells were centrifuged (5 min at 180xg) and adjusted to a cell concentration of 1 – 1.5 x10^7/mL in HBSS. Then the cells were sorted according to their fluorescence protein signal. The cells were taken up in test tubes already containing culture medium, then centrifuged and seeded into suitable culture vessels depending on the number of cells sorted.

4. ELISA

The MCP-1 titres in supernatants of stably transfected CHO-DG44 cells were quantified by ELISA with the OptEIA Human MCP-1 set according to the manufacturer's instructions (BD Biosciences Pharmingen, Heidelberg, DE). The IgG2 mAb in the supernatants from stably transfected CHO-DG44 cells was quantified by ELISA according to standard procedures (Current Protocols in Molecular Biology, Ausubel et al., 1994, updated), using on the one hand a

goat anti human IgG Fc fragment (Dianova, Hamburg, DE) and on the other hand an AP-conjugated goat anti human kappa light chain antibody (Sigma). Purified IgG2 antibody was used as the standard.

Productivities (pg/cell/day) were calculated by the formula pg/((Ct-Co) t / ln (Ct-Co)), where Co and Ct are the cell count on seeding and harvest, respectively, and t is the cultivation time.

5. Dot Assay for determining the NPT enzyme activity

In order to prepare a cell extract according to a method of Duch et al., 1990, $6x10^6$ cells were washed twice with PBS and then resuspended in 600 µL of extraction buffer (0.135 M Tris-HCl pH 6.8, 20% glycerol, 4 mM dithiothreitol). After four cycles of freezing and thawing in a bath of dry ice or water the cell debris was removed by centrifuging and the supernatant was used for the subsequent enzyme assay. The protein concentration in the cell extracts was determined by a Bradford assay using the BIO-RAD protein assay (Bio-Rad Laboratories GmbH, Munich, DE), with BSA as the standard protein (Current Protocols in Molecular Biology, Ausubel et al., 1994, updated). In order to determine the NPT enzyme activity a Dot Assay was carried out, based on the protocol of Platt et al. 1987. For this, 5 µg, 2.5 µg and 1.25 µg of protein were adjusted with extraction buffer to a final volume of 20 µL, topping up to a total protein content of 5 µg with cell extract from non-transfected CHO-DG44 cells. After the addition of 200 µL of assay buffer (67 mM Tris-HCl pH 7.1, 42 mM $MgCl_2$, 400 mM NH_4Cl) plus/minus 40 µg/mL G418 and plus/minus 5 µCi [γ-^{33}P]-ATP/mL (NEN) the extracts were incubated at 27°C for 135 minutes. Then the extracts were filtered in a 96 well vacuum manifold (Schleicher & Schüll, Dassel, DE) through a sandwich of one layer of Whatman 3MM paper, P81 phosphocellulose membrane (Whatman Laboratory Division, Maidstone, Great Britain) and nitrocellulose membrane (Schleicher & Schüll). Proteins phosphorylated by protein kinases and non-phosphorylated proteins bind to the nitrocellulose, while phosphorylated G418 passes through the nitrocellulose and binds to the phosphocellulose. After washing three times with deionised H_2O the membranes were removed from the apparatus,

washed again with H_2O and then air-dried. The radioactive signals were quantified using a Phospho Imager (Molecular Dynamics, Krefeld, DE).

6. Northern Blot Analysis

Total RNA was isolated from the cells with the TRIZOL reagent according to the manufacturer's instructions (Invitrogen GmbH, Karlsruhe, DE) and the separation of 30 µg RNA by gel electrophoresis and the transfer to a Hybond N+ nylon membrane (Amersham Biosciences, Freiburg, DE) were carried out according to the standard procedure for glyoxal/DMSO-denatured RNA (Current Protocols in Molecular Biology, Ausubel et al., 1994, updated). The probe used for the subsequent non-radioactive hybridisation with the GeneImages CDP-Star Detection Kit (Amersham Biosciences) was a PCR product which comprises the coding region of the NPT gene, FITC-dUTP-labelled according to the manufacturer's instructions with the GeneImages random prime labelling kit (Amersham Biosciences, Freiburg, DE).

7. Dot Blot Analysis

Genomic DNA was isolated from the cells using a DNA isolation kit according to the manufacturer's instructions (DNA Isolation Kit for Cells and Tissue; Roche Diagnostics GmbH, Mannheim, DE). Various amounts of DNA between 10 µg and 0.32 µg were filtered by the standard method (Ausubel et al., 1994) in an alkaline buffer using a 96 well vacuum manifold (Schleicher & Schüll, Dassel, DE) onto a Hybond N+ nylon membrane (Amersham Biosciences, Freiburg, DE). The probe used for the subsequent non-radioactive hybridisation with the GeneImages CDP-Star Detection Kit (Amersham Biosciences) was a PCR product which comprised the coding region of the NPT gene, FITC-dUTP-labelled according to the manufacturer's instructions with the GeneImages random prime labelling kit (Amersham Biosciences, Freiburg, DE).

Example 1: Mutagenesis of the neomycin-phosphotransferase

The base substitutions in the wild-type NPT-gene needed to prepare the NPT mutants Glu182Asp (SEQ ID NO:3), Asp190Gly (SEQ ID NO:5), Asp208Gly (SEQ ID NO:7) and Asp227Gly (SEQ ID NO:9) were carried out by PCR using mutagenic primers (Fig. 3). The vector pBK-CMV (Stratagene, La Jolla, USA) was used as a template for the PCR mutagenesis. First, the 5′ or 3′sections of the mutants were prepared in separate PCR operations. For the amplification primer combinations were used which consisted of Neofor5 (SEQ ID NO:11) and the relevant mutagenic reverse (rev) primer or Neorev5 (SEQ ID NO:20) and the relevant mutagenic forward (for) primer:

- In the case of NPT mutants Glu182Asp (SEQ ID NO:3) Neofor5 (SEQ ID NO:11) and E182Drev (SEQ ID NO:13) or of Neorev5 (SEQ ID NO:20) and E182Dfor (SEQ ID NO:12);
- in the case of NPT mutants Asp190Gly (SEQ ID NO:5) Neofor5 (SEQ ID NO:11) and D190Grev (SEQ ID NO:15) or of Neorev5 (SEQ ID NO:20) and D190Gfor (SEQ ID NO:14);
- in the case of NPT mutants Asp208Gly (SEQ ID NO:7) Neofor5 (SEQ ID NO:11) and D208Grev (SEQ ID NO:17) or of Neorev5 (SEQ ID NO:20) and D208Gfor (SEQ ID NO:16);
- in the case of NPT mutants Asp227Gly (SEQ ID NO:9) Neofor5 (SEQ ID NO:11) and D227Grev (SEQ ID NO:19) or of Neorev5 (SEQ ID NO:20) and D227Gfor (SEQ ID NO:18).

Then the coding strand of the 5′section and the complementary strand of the 3′section of the mutants in question were combined by hybridisation in the overlapping region formed by the mutagenic primer sequences, the single strand regions were filled in and the entire product was amplified again in a PCR with the primers Neofor5 (SEQ ID NO:11) and Neorev5 (SEQ ID NO:20). These PCR products were digested with StuI/BstBI. Then in the vector pBK-CMV part of the wild-type NPT sequence was eliminated by StuI/BstBI digestion and replaced by the corresponding fragments of the PCR products. By sequence analysis of both the complementary and the coding strand the desired base substitutions in the various mutants were verified to ensure that the remaining DNA sequence corresponded to the wild-type NPT sequence.

The NPT mutants were then isolated as a 1640 bp Bsu36I fragment, the fragment ends were filled with Klenow-DNA-polymerase and ligated with the 3750 bp Bsu36I/StuI fragment of the vector pBID, which was also treated with Klenow-DNA-polymerase. In this way the expression vectors pKS-N5, pKS-N6, pKS-N7 and pKS-N8 which contain the NPT mutants Glu182Asp, Asp190Gly, Asp208Gly or Asp227Gly were generated.
Then the human MCP-1 cDNA was cloned into these expression vectors as a 0. 3 kb HindIII/EcoRI fragment or the light chain of the humanised IgG2 antibody was cloned therein as a 0.7 kb HindIII/BamHI fragment (Fig. 4).

Example 2: Influence of the NPT mutations on the selection of stably transfected MCP-1-expressing cells

MCP-1 was chosen as an example of the expression of a single-chained protein in CHO cells. For this, CHO-DG44 was transfected with pKS-N5-MCP1, pKS-N6-MCP1, pKS-N7-MCP1, pKS-N8-MCP1 or pBIN-MCP1 (Fig. 2 and 4). Each procedure was carried out in duplicate. Two days after transfection the cells were seeded into a 96 well plate (2000 cells/well) and selected with 400 µg/mL of G418 in HT-supplemented CHO-S-SFMII medium. In the case of the cells transfected with pBIN-MCP1 , selection with 800 µg/mL G418 was carried out in parallel. The cell populations obtained were successively transferred via 24 well plates into 6 well plates. Even during the selection phase differences could be detected in the various transfection batches. In contrast to the cell populations in which selection was carried out with an NPT wild-type gene (SEQ ID NO:1), fewer cells survived the initial selection with G418 in the cell populations which had been transfected with a mutated NPT. These cell populations could not therefore be transferred into the 24 well plates until about 4 days later. And in the batches transfected with pKS-N6-MCP1 and pKS-N7-MCP1 no stably transfected cells at all could be selected at a concentration of 400 µg/mL G418. Presumably, in the NPT mutants with the mutations Asp190Gly or Asp208Gly, the enzyme function is so severely affected that it is no longer possible to inactivate enough G418 molecules to allow the stably transfected cells to grow. When the G418

concentration was lowered to 200 µg/mL a few cells survived the first selection phase but they all showed severe impairment of their growth and vitality and expansion was not possible, apart from a few exceptions in the mutant Asp208Gly.

Of the cells transfected with the mutants Glu182Asp and Asp227Gly or with the NPT-wild-type cells ,18 pools (9 pools each from batches 1 and 2) were cultivated in 6 well plates for four runs and the concentration of the MCP-1 produced was determined in the cell culture supernatant by ELISA. Cell pools in which the NPT mutants were used as selectable markers showed on average 50% – 57% (Glu182Asp mutant) or 57% – 65% (Asp227Gly mutant) higher productivities than cell pools in which the selection had been carried out with the NPT wild-type at 400 or even 800 µg/mL of G418 (Fig. 5). Thus by using NPT mutants as the selectable marker the proportion of high producers in the transfected cell populations was actually increased.

Example 3: Influence of the NPT mutations on the selection of stably transfected mAb expressing cells

In a co-transfection CHO-DG44 cells were transfected with the plasmid combination pBIDG- HC/pBIN-LC (NPT-wild-type), pBIDG-HC/pKS-N5-LC (Glu182Asp NPT mutant) or pBIDG-HC/pKS-N8-LC (Asp227Gly NPT mutant) (Fig. 2 and 4). In the vector configurations used the two protein chains of a humanised IgG2 antibody were each expressed by their own vector, which additionally also codes for a DHFR or neomycin selectable marker in a separate transcription unit.

In all, four transfection series were carried out with 6 pools per plasmid combination. In contrast to the cell populations in which selection was done with an NPT wild-type gene, fewer cells survived the initial selection with G418 in the cell populations which had been transfected with a mutated NPT. After a two- to three-week selection of the transfected cell pools in HT-free CHO-S-SFMII medium with the addition of 400 µg/mL of G418the antibody titre in the cell culture supernatants was determined by ELISA over several runs (6 – 8). In Fig. 6 the averages of the titres and productivities determined

from the 6 pools of a transfection series are shown. Compared with the use of an NPT wild-type gene as the selectable marker the cells which had been selected with the Glu182Asp mutant showed on average an increase in productivity and titre of 86% and 77%, respectively, and the cells selected with the Asp227Gly mutant even showed an increase in productivity and titre of 126% and 107%, respectively. Thus, by using an NPT mutant with reduced enzyme activity, it was possible to selectively enrich cells with a basic productivity up to twice as high.

In another transfection series the influence of different G418 concentrations on the selection was tested. To select the transfected cell pools, 3 pools in each case, 400, 500 or 600 µg/mL of G418 were used. At the higher concentration significantly fewer cells survived the initial selection in the cell populations in which selection was carried out with an NPT wild-type gene, the effect being greatest with the Asp227Gly mutant. The stably transfected cell populations obtained, however, showed no impairment of their growth or vitality. However, no significant difference could be detected between the productivities and titres obtained within a plasmid combination used for the transfection. Here again, though, the cells selected with the NPT mutants had the highest productivities on average, led by the Asp227Gly mutant with a productivity four times higher than the NPT wild-type, followed by the Glu182Asp mutant with a productivity higher by a factor of 2.4 (Fig. 7).

The vector pBIDG-HC contains another selectable marker, GFP. The GFP is transcriptionally linked to the heavy chain via an IRES element. The resulting correlation between the expression of the target protein and the selectable marker GFP therefore also makes it possible to rapidly evaluate the level and distribution of the expression levels in the transfected cell populations on the basis of the GFP fluorescence determined in FACS analyses. This is shown by way of example in Fig. 8 for the last transfection series. After two to three weeks' selection of the transfected cell pools in HT-free CHO-S-SFMII medium with the addition of G418 the GFP fluorescence was measured in a FACS analysis. The GFP fluorescence signals in fact correlated with the titre data obtained for the monoclonal IgG2 antibody. Pools selected with the NPT mutant Asp227Gly also had the higher proportion of cells with a high GFP fluorescence, followed by the cells selected with the NPT mutant Glu182Asp.

Example 4: Determining and comparing the NPT enzyme activity

In order to compare the enzyme activity of the NPT mutants with that of the NPT wild-type a Dot Assay was carried out to determine the NPT activity in cell extracts, based on the procedure of Platt et al. 1987. Cell extracts were prepared from two different mAb-expressing cell pools which had been transfected and selected either with the NPT wild-type gene (SEQ ID NO:1) or with the NPT mutants Glu182Asp (SEQ ID NO:3) or Asp227Gly (SEQ ID NO:9). . Cell extracts from untransfected CHO-DG44 cells were used as the negative control.

The enzyme activities of the NPT mutants were significantly reduced, compared with the NPT wild-type. On average the NPT mutant Asp227Gly had only 30% and the NPT mutant Glu182Asp had only 22% of the wild-type enzyme activity (Fig. 9). The signals obtained on the phosphocellulose were specific to the phosphorylation of G418 caused by the NPT enzyme activity. Without the addition of the NPT substrate G418 to the assay buffer no activity could be detected on the phosphocellulose. The signals obtained on the nitrocellulose membrane, which resulted from the proteins phosphorylated by protein kinases within the cell extract, were used as an internal control for identical amounts of sample applied.

The reduced enzyme activity of the NPT mutants compared with the NPT wild-type cannot be put down to reduced gene expression. On the contrary, Northern Blot analyses on total RNA showed that cell pools which had been transfected with the NPT wild-type and exhibited a high NPT enzyme activity expressed less RNA than cell pools transfected with NPT mutants (Fig. 10). This negative correlation between enzyme activity and the amount of mRNA reflects differences in the transcriptional activity. Transfected cells in which markers with a reduced enzyme activity were used for the selection obviously only survived the selection pressure when the exogenous DNA was integrated into transcription-active genomic regions. Only in this way are these cells able to synthesise enough marker protein to compensate for the reduced resistance to the selective agent under identical selection pressure. Dot Blot analyses carried out on genomic DNA from these transfected cell populations

ruled out the possibility that the higher expression of the NPT mutants was brought about by a gene dose effect. The hybridisation intensities of the NPT-specific signals observed in cell pools which had been transfected either with the NPT wild-type or with the NPT mutants indicated comparable copy numbers of the integrated NPT genes (Fig. 11).

Example 5: Isolation of cells with a high expression of an mAb by GFP-based FACS sorting

In a co-transfection CHO-DG44 cells were transfected with the plasmid combination pBID-HC and pBING-LC, coding for a monoclonal humanised IgG2 antibody (Fig.2 and 4). In the vector configurations used the two protein chains of the antibody are expressed by their own vector, which additionally also codes for a DHFR or modified neomycin-phosphotransferase selectable marker (Asp227Gly mutant; SEQ ID NO:9) in a separate transcription unit. In addition, another selectable marker, GFP, is contained in the vector pBING-LC. By the transcriptional linking of the expression of the GFP and the light chain by means of an IRES element, in the co-transfection of CHO-DG44 with the vectors pBID-HC/pBING-LC, cells with a high expression of the monoclonal antibody could rapidly be isolated solely by selecting the cells with a high GFP content using sequential FACS sorting. In all, 8 separate cell pools were transfected from which stably transfected cell populations were obtained after a first two to three week selection in HT-free CHO-S-SFMII medium with the addition of 400 µg/mL of G418. The titres and productivities of all 8 pools were determined by ELISA over several runs (7 – 8). Figure 12 shows the averages of these measurements. For the subsequent sequential FACS-based sorting pools 5 and 8 were selected, pool 5 having the highest productivity and pool 8 having a productivity corresponding to the average of all the pools. In each step the 5% of cells with the highest GFP fluorescence were sorted by FACS and further cultivated in the pool. This sorting was carried out up to six times in all, leaving a cultivation period of about two weeks between each sorting. Astonishingly there was found to be a good correlation between mAb productivity and GFP fluorescence (Fig. 15),

although both protein chains were expressed from their own vector and in the GFP-based FACS sorting it was only possible to select for the expression of the light chain, as a result of its transcriptional coupling with GFP. The productivities could be increased simply by FACS-based sorting to 9.5 pg/cell/day (Fig. 13). Comparable data could also be obtained by functionally linking the hamster promoter with the SV40 enhancer instead of the CMV enhancer. By means of a single subsequent MTX amplification step, starting from pools 5 and 8 from the first sorting step, by adding 100 nM of MTX to the selection medium, it was possible to increase the productivity of the pools to more than 20 pg/cell/day on average (Fig. 14). The high expression levels of the fluorescent protein had no negative effects whatsoever on cell growth and vitality. During gene amplification the growth characteristics of the cells were also critically negatively affected by the addition of MTX, particularly when added in higher concentrations. The presorted cell pools did, however, exhibit considerably more robust behaviour in the presence of the very high initial dose of 100 nM of MTX. They survived the selection phase much better, i.e. cell populations with high vitality and a good growth rate were obtained after only 2 weeks.

In addition, the development time for selecting high-producing cells in comparison with a conventional stepwise gene amplification strategy, which generally comprises four amplification stages with increasing amounts of MTX, could be reduced to about 6 weeks. This was achieved by the combined use of a concentration of transfected cells with increased expression of the gene of interest, obtained by the use of a modified NPT selectable marker with reduced enzyme activity, followed by a GFP-based FACS sorting with a subsequent gene amplification step.

LITERATURE

Adam, M.A. et al., *J Virol* 1991, 65, 4985 - 4990

Ausubel, F.M. et al., Current Protocols in molecular biology. New York : Greene Publishing Associates and Wiley-Interscience. 1994 (updated)

Blázques, J. et al., *Molecular Microbiology* 1991, 5(6), 1511 – 1518

Bennett, R.P. et al., *BioTechniques* 1998, 24, 478 - 482

Burk, D.L. et al., *Biochemistry* 2001, 40 (30), 8756 - 8764

Chalfie, M. et al., *Science* 1994, 263, 802 - 805

Chamov, S.M. et al., Antibody Fusion Proteins, Wiley-Liss Inc., 1999

Davies, M.V. et al., *J Virol* 1992, 66, 1924 - 1932

Faisst, S. et al., *Nucleic Acids Research* 1992, 20, 3 – 26

Gossen, M. et al., *Curr Opi Biotech* 1994, 5, 516 - 520

Duch, M. et al., *Gene* 1990, 95, 285 - 288

Hanson, K.D. et al., *Mol Cell Biol* 1995, 15(1), 45 – 51

Haber, D.A. et al., *Somatic Cell Genetics* 1982, 8, 499 - 508

Harris et al., Protein Purification : A Practical Approach, Pickwood and Hames, eds., IRL Press, 1995

Hemann, C. et al., *DNA Cell Biol* 1994, 13 (4), 437 - 445

Hon, W. et al., *Cell* 1997, 89, 887 - 895

Hu, S. et al., *Cancer Res*. 1996, 56 (13), 3055 - 3061

Huston, C. et al., *Proc Natl Acad Sci USA* 1988, 85 (16), 5879 - 5883

Jang, S.K. et al., *J Virol* 1989, 63, 1651 – 1660

Kaufman, R.J., *Methods in Enzymology* 1990, 185, 537 - 566

Kocabiyik, S. et al., *Biochem Biophys Res Commun* 1992, 185(3), 925 – 931

Kortt, A.A. et al., *Protein Engineering* 1997, 10 (4), 423 - 433

Lottspeich F. and Zorbas H. eds., Bioanalytic, Spektrum Akad. Verl., 1998

Lovejoy, B. et al., *Science* 1993, 259, 1288 - 1293

Monaco, L. et al., *Gene* 1996, 180, 145 – 15

Morgan, R.A. et al., *Nucleic Acids Research* 1992, 20, 1293 - 1299

Mosser, D.D. et al., *BioTechniques* 1997, 22, 150 - 161

Pack, P. et al., *Biotechnology* 1993, 11, 1271 – 1277

Pack, P. et al., *J Mol Biol* 1995, 246 (11), 28 - 34

Pelletier, J. et al., *Nature* 1988, 334, 320 – 325

Perisic, O. et al., *Structure* 1994, 2, 1217 - 1226

Platt, S.G. et al., *Analyt Biochem* 1987, 162, 529 - 535

Ramesh, N. et al., *Nucleic Acids Research* 1996, 24, 2697 – 2700

Sambrook, J., Fritsch, E.F. & Maniatis, T.,Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989

Scopes, R., Protein Purification, Springer Verlag, 1988

Shaw, K.J. et al., *Microbiological Reviews* 1993, 57(1), 138 - 163

Simonson, C.C. et al., *Proc Natl Acad Sci USA* 1983, 80, 2495 - 2499

Sugimoto et al., *Biotechnology* 1994, 12, 694 - 698

Yenofsky, R.L. et al., *Proc Natl Acad Sci USA* 1990, 87, 3435 – 3439

Yoshimura, T. et al., *FEBS LETTERS* 1989, 244(2), 487 - 493

Urlaub, G. et al., *Cell* 1983, 33, 405 – 412

Werner, R.G. et al., *Arzneim.-Forsch./Drug.Res.* **1998**, 48, 870 - 880

Wigler, M. et al., *Proc Natl Acad Sci USA* **1980**, 77, 3567 - 3570

SEQUENCE LISTING

<110> BOEHRINGER INGELHEIM PHARMA KG

<120> Eukaryotic expression vector comprising a modified neomycin-phosphotransferase gene, and methods for the selection of recombinant cells producing high levels of a desired gene product

<130> Case 1-1411

<140>
<141>

<160> 21

<170> PatentIn Ver. 2.1

<210> 1
<211> 795
<212> DNA
<213> Escherichia coli

<400> 1
atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc 60
ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca 120
gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 180
caagacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg 240
ctcgacgttg tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag 300
gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg 360
cggcggctgc atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc 420
atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa 480
gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgag catgcccgac 540
ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat 600
ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac 660
atagcgttgg ctacccgtga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc 720
ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt 780
gacgagttct tctga 795

<210> 2
<211> 264
<212> PRT
<213> Escherichia coli

<400> 2
Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val
1 5 10 15

Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
20 25 30

Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
35 40 45

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
50 55 60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val
65 70 75 80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu
85 90 95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
100 105 110

Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
115 120 125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
130 135 140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu
145 150 155 160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala
165 170 175

Ser Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys
180 185 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp
195 200 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala
210 215 220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe
225 230 235 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe
245 250 255

Tyr Arg Leu Leu Asp Glu Phe Phe
260

<210> 3
<211> 795
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Neomycin mutant E182D

<400> 3
atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc 60
ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca 120
gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 180
caagacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg 240
ctcgacgttg tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag 300
gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg 360
cggcggctgc atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc 420
atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa 480
gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgag catgcccgac 540
ggcgatgatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat 600
ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac 660
atagcgttgg ctacccgtga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc 720
ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt 780

gacgagttct tctga 795

<210> 4
<211> 264
<212> PRT
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Neomycin mutant E182D

<400> 4
Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val
1 5 10 15

Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
20 25 30

Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
35 40 45

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
50 55 60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val
65 70 75 80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu
85 90 95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
100 105 110

Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
115 120 125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
130 135 140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu
145 150 155 160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala
165 170 175

Ser Met Pro Asp Gly Asp Asp Leu Val Val Thr His Gly Asp Ala Cys
180 185 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp
195 200 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala
210 215 220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe
225 230 235 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe
245 250 255

Tyr Arg Leu Leu Asp Glu Phe Phe
260

<210> 5
<211> 795
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Neomycin mutant D182G

<400> 5
atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc 60
ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca 120
gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 180
caagacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg 240
ctcgacgttg tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag 300
gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg 360
cggcggctgc atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc 420
atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa 480
gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgag catgcccgac 540
ggcgaggatc tcgtcgtgac ccatggcggt gcctgcttgc cgaatatcat ggtggaaaat 600
ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac 660
atagcgttgg ctacccgtga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc 720
ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt 780
gacgagttct tctga 795

<210> 6
<211> 264
<212> PRT
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Neomycin mutant D182

<400> 6
Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val
1 5 10 15

Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
20 25 30

Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
35 40 45

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
50 55 60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val
65 70 75 80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu
85 90 95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
100 105 110

```
Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
        115                 120                 125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
    130                 135                 140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu
145                 150                 155                 160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala
                165                 170                 175

Ser Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Gly Ala Cys
            180                 185                 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp
        195                 200                 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala
    210                 215                 220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe
225                 230                 235                 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe
                245                 250                 255

Tyr Arg Leu Leu Asp Glu Phe Phe
            260


<210> 7
<211> 795
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
      Neomycin mutant D208G

<400> 7
atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc 60
ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca 120
gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 180
caagacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg 240
ctcgacgttg tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag 300
gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg 360
cggcggctgc atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc 420
atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa 480
gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgag catgcccgac 540
ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat 600
ggccgctttt ctggattcat cggctgtggc cggctgggtg tggcggaccg ctatcaggac 660
atagcgttgg ctacccgtga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc 720
ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt 780
gacgagttct tctga                                                  795


<210> 8
<211> 264
<212> PRT
<213> Artificial sequence
```

<220>
<223> Description of the artificial sequence:
Neomycin mutant D208G

<400> 8
Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val
1 5 10 15

Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
20 25 30

Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
35 40 45

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
50 55 60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val
65 70 75 80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu
85 90 95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
100 105 110

Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
115 120 125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
130 135 140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu
145 150 155 160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala
165 170 175

Ser Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys
180 185 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Gly
195 200 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala
210 215 220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe
225 230 235 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe
245 250 255

Tyr Arg Leu Leu Asp Glu Phe Phe
260

<210> 9
<211> 795
<212> DNA

<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Neomycin mutant D227G

<400> 9
atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc 60
ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca 120
gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 180
caagacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg 240
ctcgacgttg tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag 300
gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg 360
cggcggctgc atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc 420
atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa 480
gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgag catgcccgac 540
ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat 600
ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac 660
atagcgttgg ctacccgtgg tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc 720
ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt 780
gacgagttct tctga 795

<210> 10
<211> 264
<212> PRT
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Neomycin mutant D227G

<400> 10
Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val
1 5 10 15

Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
20 25 30

Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
35 40 45

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
50 55 60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val
65 70 75 80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu
85 90 95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
100 105 110

Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
115 120 125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
130 135 140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu

145 150 155 160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala
165 170 175

Ser Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys
180 185 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp
195 200 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala
210 215 220

Thr Arg Gly Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe
225 230 235 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe
245 250 255

Tyr Arg Leu Leu Asp Glu Phe Phe
260

<210> 11
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Oligonucleotide Neofor5

<400> 11
ttccagaagt agtgaggagg c 21

<210> 12
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Oligonucleotide E182Dfor

<400> 12
gacggcgatg atctcgtcgt 20

<210> 13
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Oligonucleotide E182Drev

<400> 13
acgacgagat catcgccgtc 20

<210> 14
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Oligonucleotide D190Gfor

<400> 14
catggcggtg cctgcttgc 19

<210> 15
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Oligonucleotide D190Grev

<400> 15
gcaagcaggc accgccatg 19

<210> 16
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Oligonucleotide D208Gfor

<400> 16
gattcatcgg ctgtggccg 19

<210> 17
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Oligonucleotide D208Grev

<400> 17
cggccacagc cgatgaatc 19

<210> 18
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:

Oligonucleotide D227Gfor

<400> 18
ctacccgtgg tattgctgaa g 21

<210> 19
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Oligonucleotide D227Grev

<400> 19
cttcagcaat accacgggta g 21

<210> 20
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
Oligonucleotide Neorev5

<400> 20
atggcaggtt gggcgtcgc 19

<210> 21
<211> 2406
<212> DNA
<213> Cricetulus griseus

<300>
<310> PCT/EP/96/04631
<311> 1996-10-24
<312> 1997-05-01

<400> 21
gatctccagg acagccatgg ctattacaca gagaaaccct gtctggaaaa acaaaaaatt 60
agtgtccatg tgtaaatgtg tggagtatgc ttgtcatgcc acatacagag gtagagggca 120
gtttatggga gtcagttcct attcttcctt tatgggggac ctggggactg aactcaggtc 180
atcaggcttg gcagaaagtg cattagctca cggagcctta tcattggcga aagctctctc 240
aagtagaaaa tcaatgtgtt tgctcatagt gcaatcatta tgtttcgaga ggggaagggt 300
acaatcgttg gggcatgtgt ggtcacatct gaatagcagt agctccctag gagaattcca 360
agttctttgg tggtgtatca atgcccttaa aggggtcaac aacttttttt ccctctgaca 420
aaactatctt cttatgtcct tgtccctcat atttgaagta ttttattctt tgcagtgttg 480
aatatcaatt ctagcacctc agacatgtta ggtaagtacc ctacaactca ggttaactaa 540
tttaatttaa ctaatttaac cccaacactt tttctttgtt tatccacatt tgtggagtgt 600
gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgc 660
gcgcgcgcgc gcgctcggat cattctacct tttgtttaaa aaatgttagt ccaggggtgg 720
ggtgcactgt gaaagtctga gggtaacttg ctggggtcag ttctttccac tataggacag 780
aactccaggt gtcaactctt tactgacaga accatccaaa tagccctatc taattttagt 840
tttttattta tttatttttt gtttttcgag acaggtttc tctgtggctt tggaggctgt 900
cctggaacta gctcttgtag accaggctgg tctcgaactc agagatccac ctgcctctgc 960
ctcctgagtg ctgggattaa aggcatgcgc caccaacgct tggctctacc taattttaaa 1020
agagattgtg tgtcacaagg gtgtcatgtc gccctgcaac caccccccc ccaaaaaaa 1080

```
aaaaaaaaaa acttcactga agctgaagca cgatgatttg gttactctgg ctggccaatg 1140
agctctaggg agtctcctgt caaacagaat ctcaacaggc gcagcagtct tttttaaagt 1200
ggggttacaa cacaggtttt tgcatatcag gcattttatc taagctattt cccagccaaa 1260
aatgtgtatt ttggaggcag cagagctaat agattaaaat gagggaagag cccacacagg 1320
ttatttaggaa gataagcatc ttctttatat aaaacaaaac caaaccaaac tggaggaggt 1380
ctacctttag ggatggaaga aaagacattt agagggtgca atagaaaggg cactgagttt 1440
gtgaggtgga ggactgggag agggcgcaac cgctttaact gtcctgtttt gcctattttt 1500
tggggacagc acatgttcct atttttccca ggatgggcaa tctccacgtc caaacttgcg 1560
gtcgaggact acagtcattt tgcaggtttc cttactgtat ggcttttaaa acgtgcaaag 1620
gtgaccatta accgtttcac gctgggaggg cacgtgcggc tcagatgctt cctctgactg 1680
agggccagga gggggctaca cggaagaggc cacacccgca cttgggaaga ctcgatttgg 1740
gcttcagctg gctgagacgc cccagcaggc tcctcggcta caccttcagc cccgaatgcc 1800
ttccggccca taacccttcc cttctaggca tttccggcga ggacccaccc tcgcgccaaa 1860
cattcggccc catccccccgg tcctcacctg aatctctaac tctgactcca gagtttagag 1920
actataacca gatagcccgg atgtgtggaa ctgcatcttg ggacgagtag ttttagcaaa 1980
aagaaagcga cgaaaaacta caattcccag acagacttgt gttacctctc ttctcatgct 2040
aaacaagccc cctttaaagg aaagcccctc ttagtcgcat cgactgtgta agaaaggcgt 2100
ttgaaacatt ttaatgttgg gcacaccgtt tcgaggaccg aaatgagaaa gagcataggg 2160
aaacggagcg cccgagctag tctggcactg cgttagacag ccgcggtcgt tgcagcgggc 2220
aggcacttgc gtggacgcct aaggggcggg tctttcggcc gggaagcccc gttggtccgc 2280
gcggctcttc ctttccgatc cgccatccgt ggtgagtgtg tgctgcgggc tgccgctccg 2340
gcttggggct tcccgcgtcg ctctcaccct ggtcggcggc tctaatccgt ctcttttcga 2400
atgtag                                                            2406
```

PATENT CLAIMS

1. Eukaryotic expression vector, containing a heterologous gene of interest functionally linked to a heterologous promoter and a modified neomycin-phosphotransferase gene, which codes for a neomycin-phosphotransferase which has a lower enzyme activity compared with wild-type neomycin-phosphotransferase.

2. Eukaryotic expression vector, containing a multiple cloning site for the incorporation of a gene which codes for a protein/product of interest and is/will be functionally linked to a heterologous promoter, and a modified neomycin-phosphotransferase gene which codes for a neomycin-phosphotransferase which has a lower enzyme activity compared with wild-type neomycin-phosphotransferase.

3. Expression vector according to claim 1 or 2, characterised in that the modified neomycin-phosphotransferase gene is a mutant which codes for a different amino acid at amino acid position 182 or 227 than the wild-type gene.

4. Expression vector according to claim 3, characterised in that the neomycin-phosphotransferase gene is the mutant Glu182Asp or Asp227Gly.

5. Expression vector according to claim 4, characterised in that it contains one or more enhancers functionally linked to the promoter or promoters.

6. Expression vector according to one of Claims 1 to 5, characterised in that it additionally contains a gene for a fluorescent protein which is/will be functionally linked to the gene of interest and the heterologous promoter.

7. Expression vector according to claim 6, characterised in that it additionally contains an internal ribosome entry site (IRES) which enables bicistronic

expression of the gene which codes for a fluorescent protein, and of the gene which codes for a protein/product of interest, under the control of a heterologous promoter.

8. Expression vector according to claim 6 or 7, characterised in that the gene which codes for a fluorescent protein, and the modified neomycin-phosphotransferase gene are located in one or in two separate transcription units.

9. Expression vector according to one of Claims 5 to 8, characterised in that the enhancer is a CMV or SV40 enhancer.

10. Expression vector according to one of Claims 1 to 9, characterised in that it contains a hamster ubiquitin/S27a promoter.

11. Expression vector according to one of Claims 1 to 10, characterised in that the heterologous gene of interest is under the control of the ubiquitin/S27a promoter.

12. Mammalian cell which has been transfected with an expression vector according to one of Claims 1 to 11.

13. Mammalian cell according to claim 12, characterised in that it has additionally been transfected with a gene for an amplifiable selectable marker.

14. Mammalian cell according to claim 13, characterised in that the amplifiable selectable marker gene is dihydrofolate-reductase (DHFR).

15. Mammalian cell according to one of Claims 12 to 14, characterised in that it has additionally been transfected with one or more vectors with genes which additionally code for one or more protein/products or sub-units of a protein of interest.

16. Mammalian cell according to one of claims 12 to 15, wherein the mammalian cell is a rodent cell.

17. Mammalian cell according to claim 16, wherein the rodent cell is a CHO or BHK cell.

18. Method of enriching mammalian cells, characterised in that
    (i) a pool of mammalian cells is transfected with a gene for a modified neomycin-phosphotransferase, which by comparison with wild-type neomycin-phosphotransferase has only 10 to 50% of the activity;
    (ii) the mammalian cells are cultivated under conditions which allow expression of the modified neomycin-phosphotransferase gene; and
    (iii) the mammalian cells are cultivated in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells, and gives preference to the growth of those cells which express the modified neomycin-phosphotransferase gene.

19. Method of obtaining and selecting mammalian cells which express at least one gene of interest, characterised in that
    (i) a pool of mammalian cells is transfected with at least one gene of interest and a gene for a modified neomycin-phosphotransferase which by comparison with wild-type neomycin-phosphotransferase has only 10 to 50% of the activity;
    (ii) the mammalian cells are cultivated under conditions which allow expression of the gene or genes of interest and of the modified neomycin-phosphotransferase gene; and
    (iii) the mammalian cells are cultivated in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells, and gives preference to the growth of those cells which express the modified neomycin-phosphotransferase gene.

20. Method according to claim 19, characterised in that said pool of mammalian cells is transfected with an expression vector according to claim 1 to 11.

21. Method of obtaining and selecting mammalian cells which express at least one heterologous gene of interest, characterised in that
(i) recombinant mammalian cells are transformed with an expression vector according to one of Claims 6 to 11;
(ii) are cultivated under conditions which allow expression of the gene (or genes) of interest, of the gene which codes for a fluorescent protein, and of the modified neomycin-phosphotransferase gene;
(iii) the mammalian cells are cultivated in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells, and gives preference to the growth of those cells which express the modified neomycin-phosphotransferase gene; and
(iv) the mammalian cells are sorted by flow-cytometric analysis.

22. Method according to one of Claims 19 to 21, characterised in that
(i) the transfected mammalian cells are additionally transfected with an amplifiable selectable marker gene;
(ii) are cultivated under conditions in which the amplifiable selectable marker gene is also expressed; and
(iii) a selecting agent which brings about amplification of the amplifiable selectable marker gene is added to the culture medium.

23. Method according to claim 22, characterised in that the amplifiable selectable marker gene is a DHFR gene and the selecting agent is methotrexate.

24. Method according to one of Claims 19 to 21, characterised in that the sorted mammalian cells exhibit an average specific productivity of more than 5pg of the desired gene product (or products) per day and per cell.

25. Method according to claim 22 or 23, characterised in that the sorted host cells exhibit an average specific productivity of more than 20pg of the desired gene product (or products) per day and per cell.

26. Method of producing at least one protein of interest in recombinant mammalian cells, characterised in that
(i) a pool of mammalian cells is transfected with at least one gene of interest and one gene for a modified neomycin-phosphotransferase which by comparison with wild-type neomycin-phosphotransferase has only 10 to 50% of the activity;
(ii) the cells are cultivated under conditions which allow expression of the gene (or genes) of interest and of the modified neomycin-phosphotransferase;
(iii) the mammalian cells are cultivated in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells, and gives preference to the growth of those cells which express the modified neomycin-phosphotransferase gene; and
(iv) the protein(s) of interest is or are obtained from the mammalian cells or the culture supernatant.

27. Method according to claim 26, characterised in that said pool of mammalian cells is transfected with an expression vector according to claim 1 to 11.

28. Method of producing at least one protein of interest, characterised in that
(i) mammalian cells according to one of Claims 12 to 17 are cultivated under conditions which allow expression of the gene of interest and of the modified neomycin-phosphotransferase gene;
(ii) the mammalian cells are cultivated in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells, and gives preference to the growth of those cells which express the modified neomycin-phosphotransferase gene; and
(iii) the protein(s) of interest is or are obtained from the mammalian cells or the culture supernatant.

29. Method of producing at least one protein of interest, characterised in that
(i) mammalian cells according to one of Claims 13 to 17 are cultivated under conditions which allow expression of the gene of interest, of the

modified neomycin-phosphotransferase gene and of the amplifiable selectable marker gene;

(ii) the mammalian cells are cultivated and selected in the presence of at least one selecting agent which acts selectively on the growth of mammalian cells, and gives preference to the growth of those cells which express the modified neomycin-phosphotransferase gene;

(iii) the selected mammalian cells are subjected to at least one gene amplification step; and

(iv) the protein(s) of interest is or are subsequently obtained from the mammalian cells or the culture supernatant.

30. Method according to one of Claims 26 to 29, characterised in that the mammalian cells are transfected with at least two genes of interest which code for a heteromeric protein/product, and

(i) are cultivated under conditions which allow expression of the subunits of the heteromeric protein/product; and

(ii) the heteromeric protein/product is isolated from the culture or culture medium.

31. Method according to one of Claims 18 to 30, characterised in that the mammalian cells are cultivated in suspension culture.

32. Method according to one of Claims 18 to 31, characterised in that the mammalian cells are cultivated serum-free.

33. Method according to one of claims 18 to 32, characterised in that the modified neomycin-phosphotransferase gene is a mutant which codes for a different amino acid than the wild-type gene at amino acid position 182 or 227.

34. Method according to one of claims 18 to 33, characterised in that by comparison with the wild-type gene the modified neomycin-phosphotransferase gene codes for aspartic acid at amino acid position 182 or for glycine at amino acid position 227.

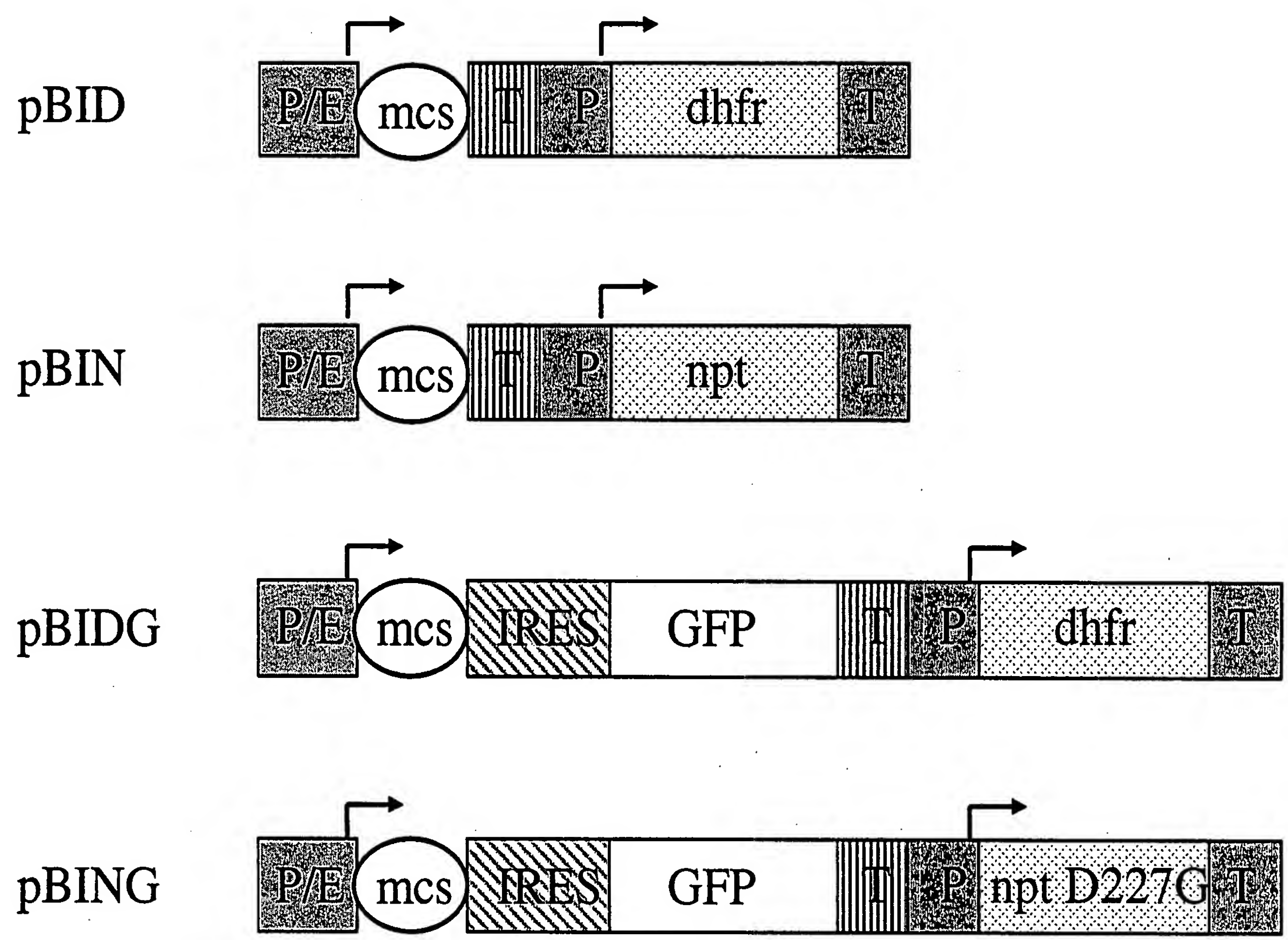
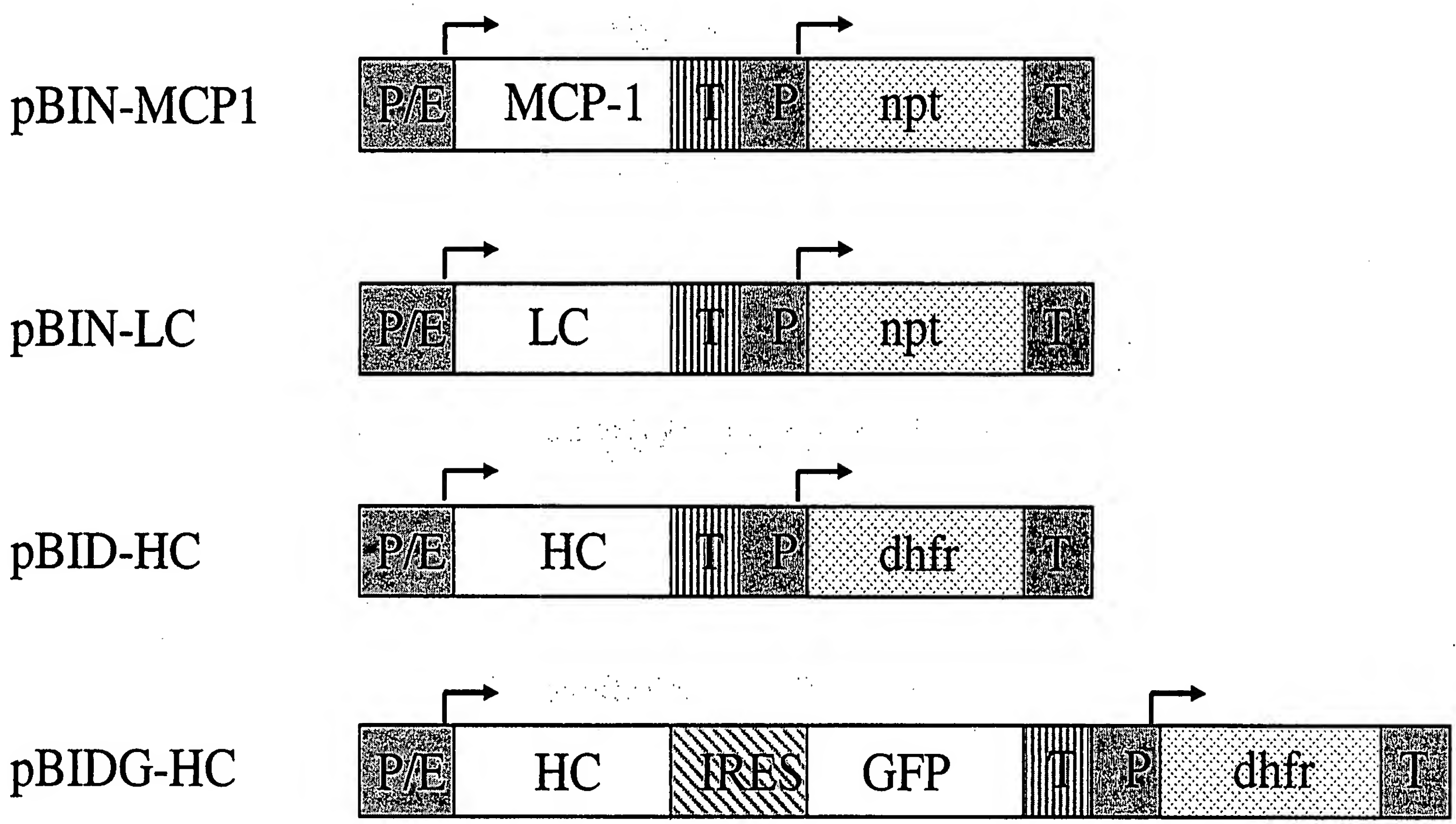

pBID
P/E
mcs
T
P
dhfr
T
pBIN
P/E
mcs
T
P
npt
T
pBIDG
P/E
mcs
IRES
GFP
T
P
dhfr
T
pBING
P/E
mcs
IRES
GFP
T
P
npt D227G
T

Figure 1



pBIN-MCP1
P/E
MCP-1
T
P
npt
T
pBIN-LC
P/E
LC
T
P
npt
T
pBID-HC
P/E
HC
T
P
dhfr
T
pBIDG-HC
P/E
HC
IRES
GFP
T
P
dhfr
T

Figure 2

```
                                                                  182                       190
                                                      Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys Leu Pro Asn Ile

              Neofor5                                             T  E182Dfor                       G  D190Gfor
5'-gagctattccagaagtagtgaggaggctttt.....CCC GAC GGC GAG GAT CTC GTC GTG ACC CAT GGC GAT GCC TGC TTG CCG AAT ATC
3'-ctcgataaggtcttcatcactcctccgaaaa.....GGG CTG CCG CTC CTA GAG CAG CAC TGG GTA CCG CTA CGG ACG AAC GGC TTA TAG
                                                                  A  E182Drev                       C  D190Grev

                                                         208
Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu

                                                         G  D208Gfor
ATG GTG GAA AAT GGC CGC TTT TCT GGA TTC ATC GAC TGT GGC CGG CTG GGT GTG GCG GAC CGC TAT CAG GAC ATA GCG TTG
TAC CAC CTT TTA CCG GCG AAA AGA CCT AAG TAG CTG ACA CCG GCC GAC CCA CAC CGC CTG GCG ATA GTC CTG TAT CGC AAC
                                                         C  D208Grev

         227
Ala Thr Arg Asp Ile Ala Glu Glu

         G  D227Gfor
GCT ACC CGT GAT ATT GCT GAA GAG....caagcgacgcccaacctgccatca-3'
CGA TGG GCA CTA TAA CGA CTT CTC....gttcgctgcgggttggacggtagt-5'
         C  D227Grev                     Neorev5
```

NPT-Mutant Glu182Asp: GAG => GAT
NPT-Mutant Asp190Gly: GAT => GGT
NPT-Mutant Asp208Gly: GAC => GGC
NPT-Mutant Asp227Gly: GAT => GGT

Figure 3

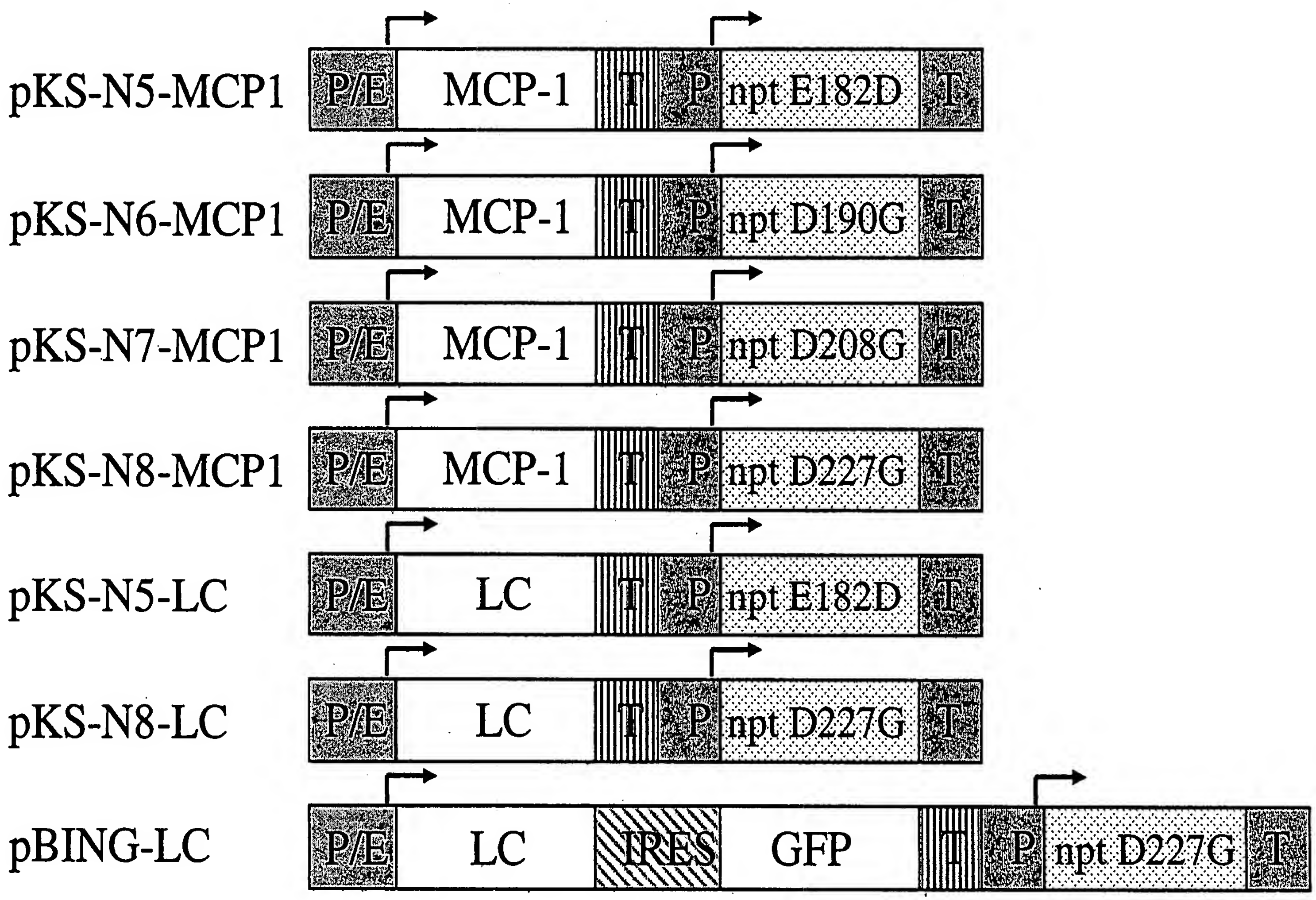
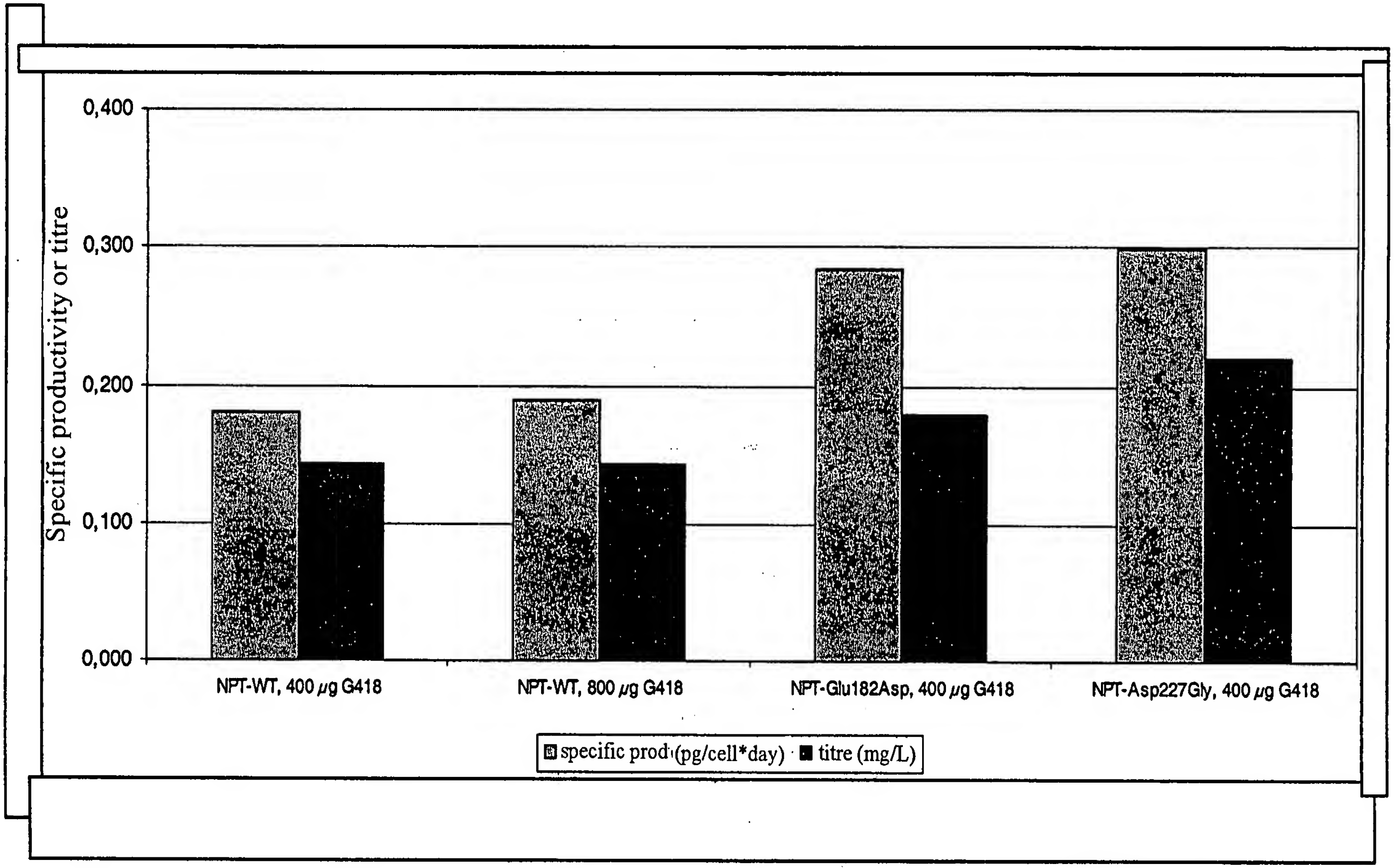

pKS-N5-MCP1
P/E
MCP-1
T
P
npt E182D
T
pKS-N6-MCP1
P/E
MCP-1
T
P
npt D190G
T
pKS-N7-MCP1
P/E
MCP-1
T
P
npt D208G
T
pKS-N8-MCP1
P/E
MCP-1
T
P
npt D227G
T
pKS-N5-LC
P/E
LC
T
P
npt E182D
T
pKS-N8-LC
P/E
LC
T
P
npt D227G
T
pBING-LC
P/E
LC
IRES
GFP
T
P
npt D227G
T

Figure 4



Specific productivity or titre
0,400
0,300
0,200
0,100
0,000
NPT-WT, 400 µg G418
NPT-WT, 800 µg G418
NPT-Glu182Asp, 400 µg G418
NPT-Asp227Gly, 400 µg G418
specific prod (pg/cell*day)
titre (mg/L)

Figure 5

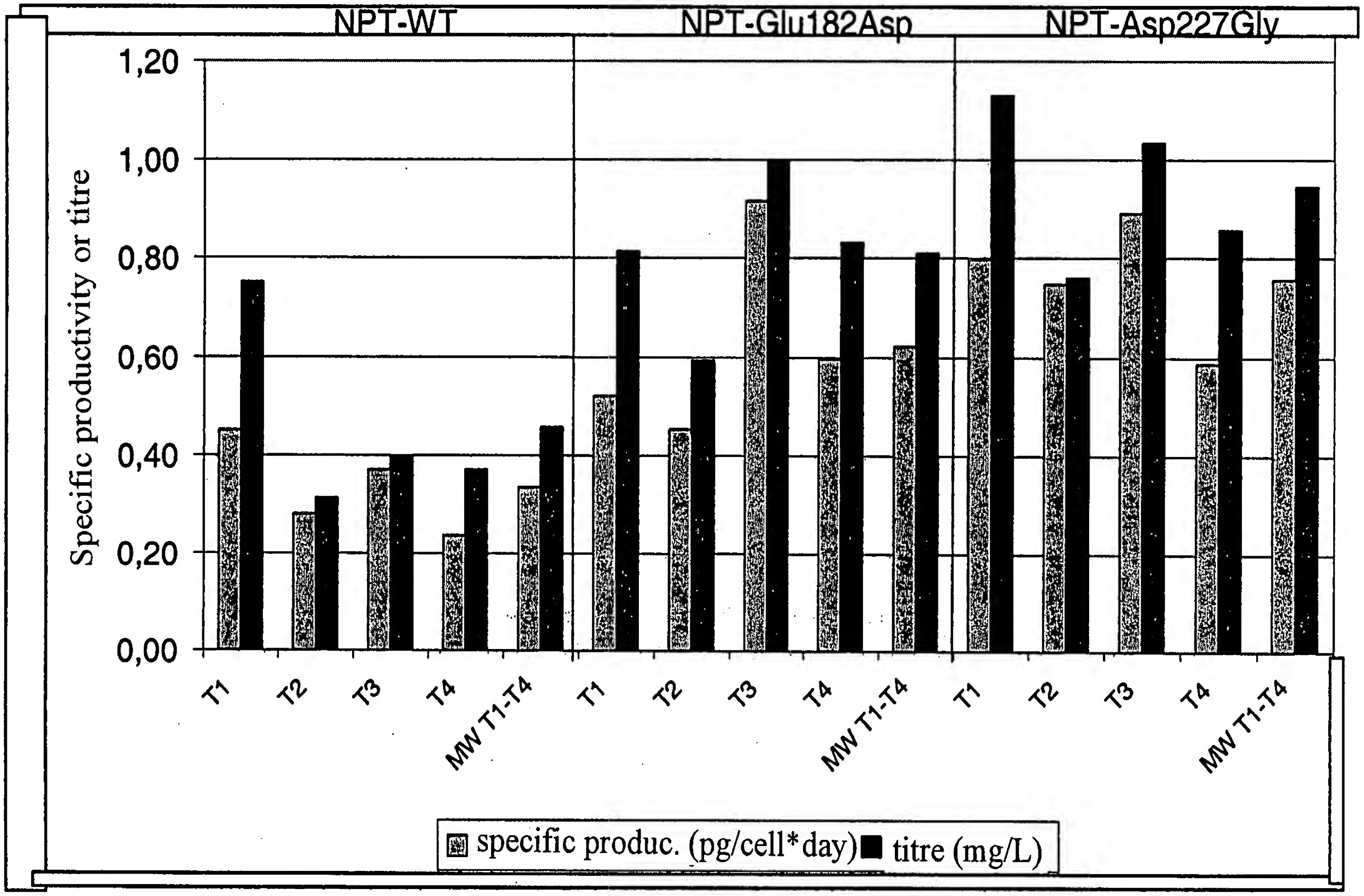
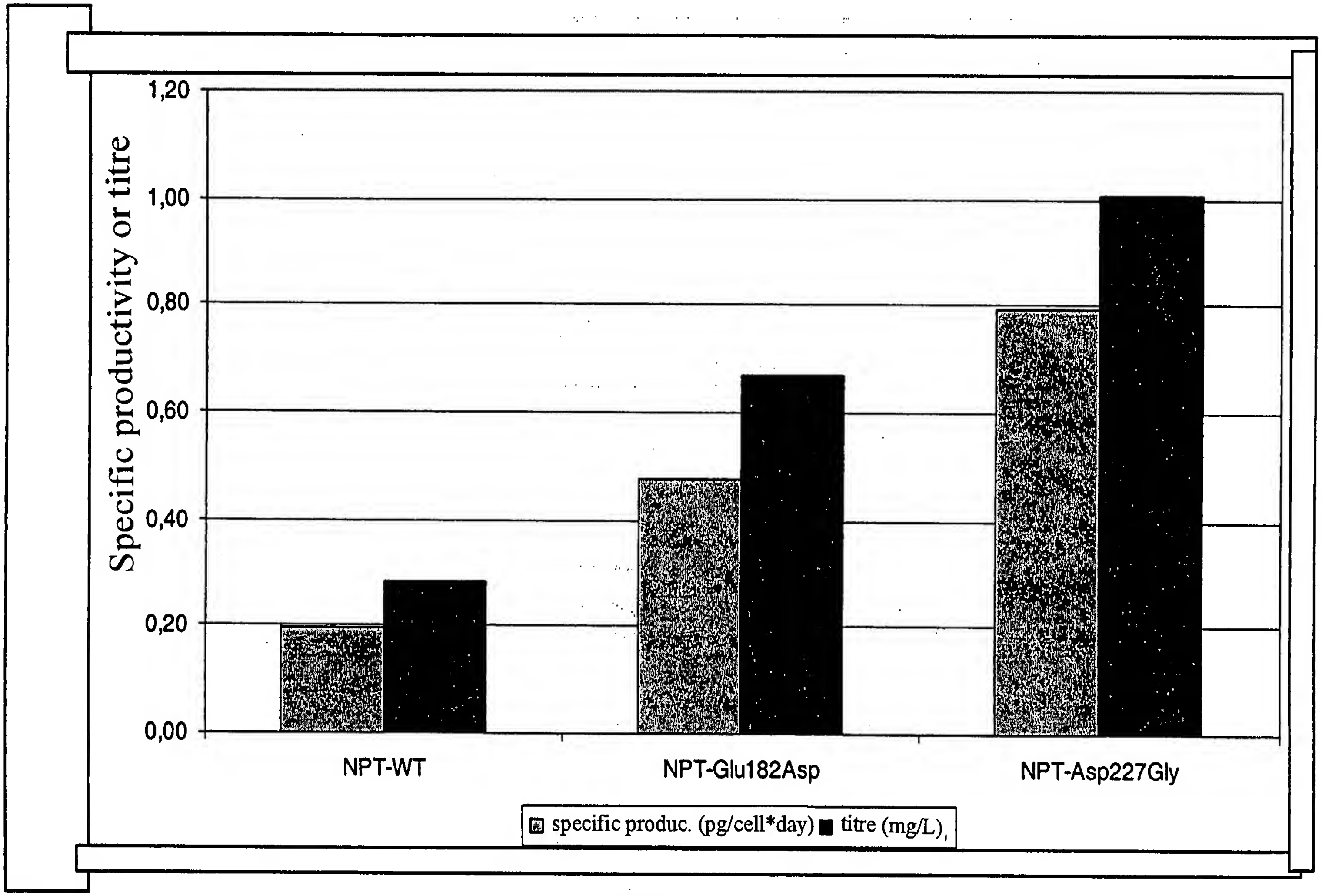

NPT-WT
NPT-Glu182Asp
NPT-Asp227Gly
1,20
1,00
0,80
0,60
0,40
0,20
0,00
Specific productivity or titre
T1
T2
T3
T4
MW T1-T4
specific produc. (pg/cell*day)
titre (mg/L)

Figure 6



1,20
1,00
0,80
0,60
0,40
0,20
0,00
Specific productivity or titre
NPT-WT
NPT-Glu182Asp
NPT-Asp227Gly
specific produc. (pg/cell*day)
titre (mg/L)

Figure 7

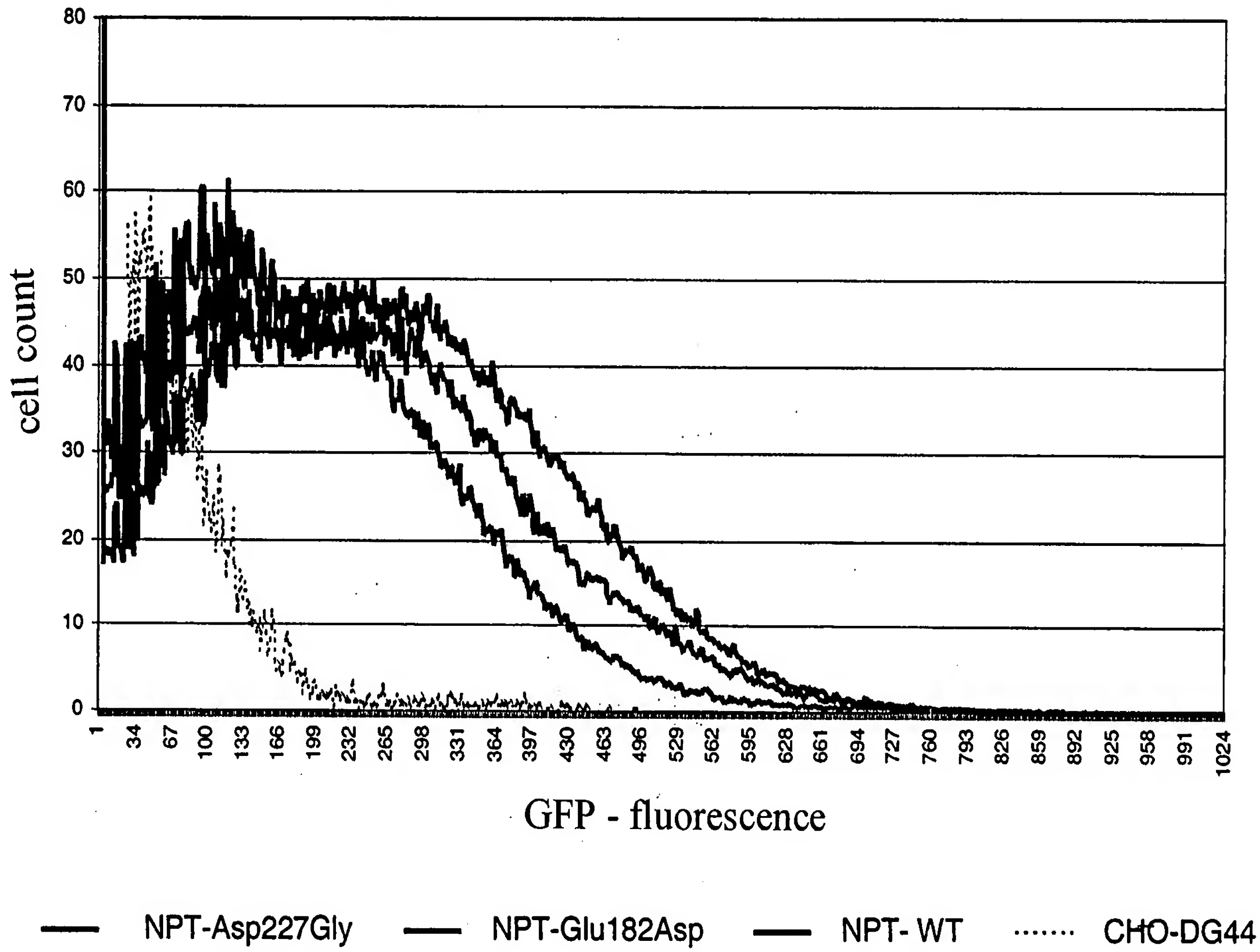
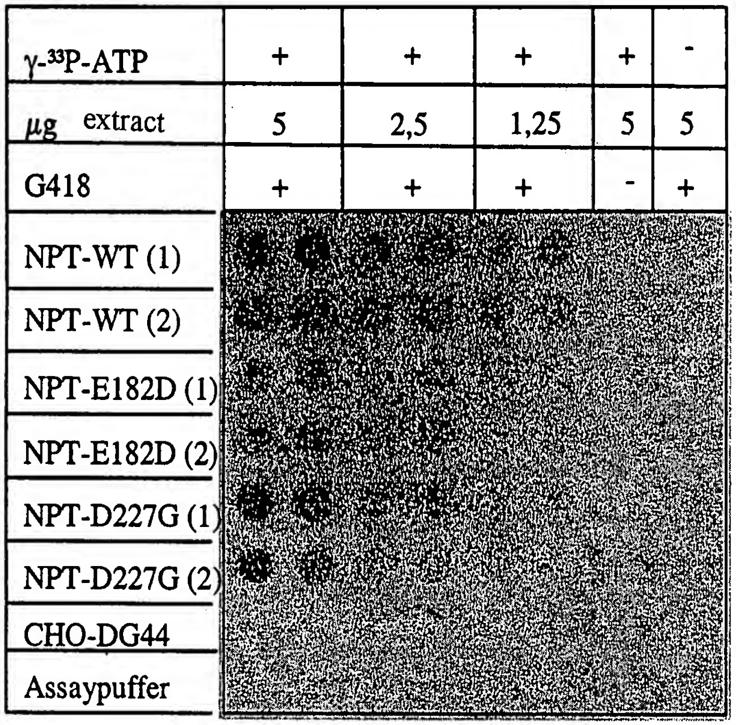
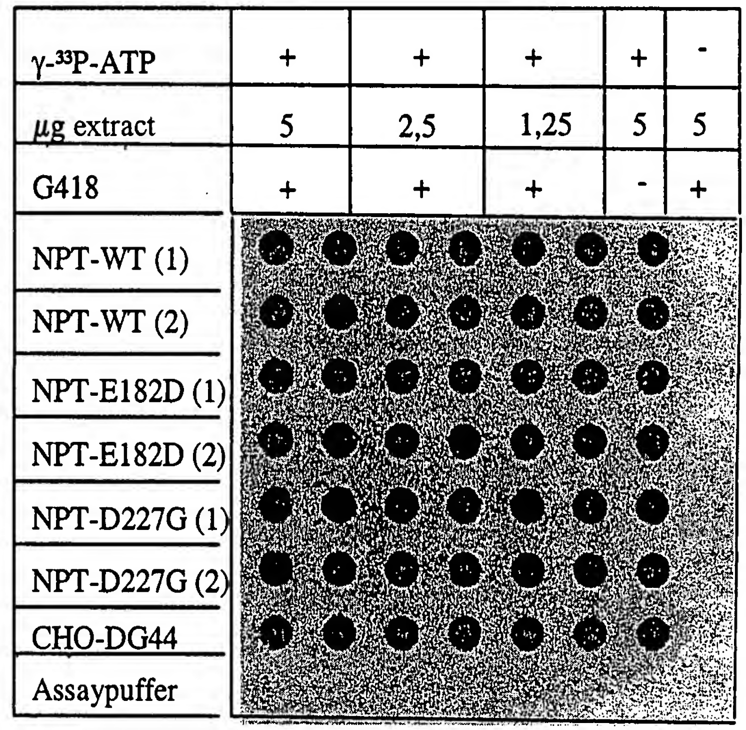

80
70
60
50
40
30
20
10
0
cell count
1
34
67
100
133
166
199
232
265
298
331
364
397
430
463
496
529
562
595
628
661
694
727
760
793
826
859
892
925
958
991
1024
GFP - fluorescence
NPT-Asp227Gly
NPT-Glu182Asp
NPT- WT
CHO-DG44

Figure 8



Phosphocellulose



Nitrocellulose

Figure 9

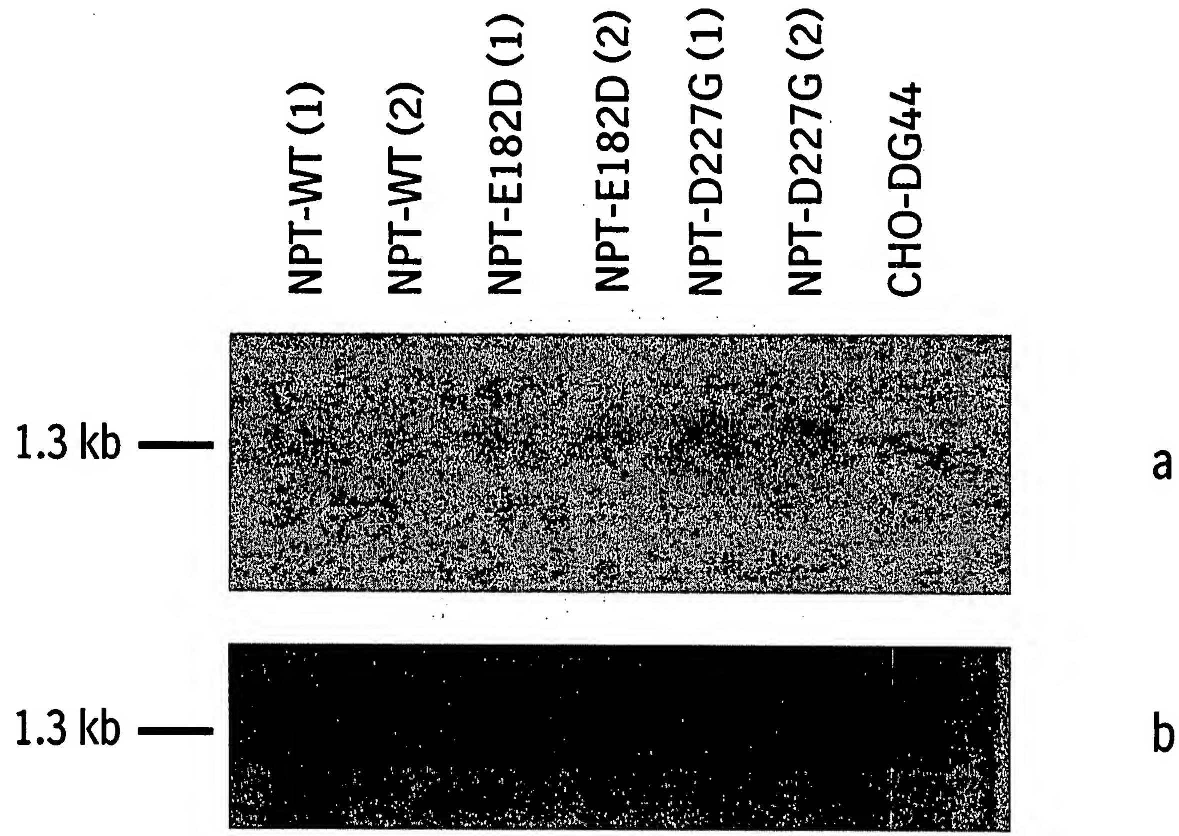

NPT-WT (1)
NPT-WT (2)
NPT-E182D (1)
NPT-E182D (2)
NPT-D227G (1)
NPT-D227G (2)
CHO-DG44
1.3 kb
a
1.3 kb
b

Figure 10

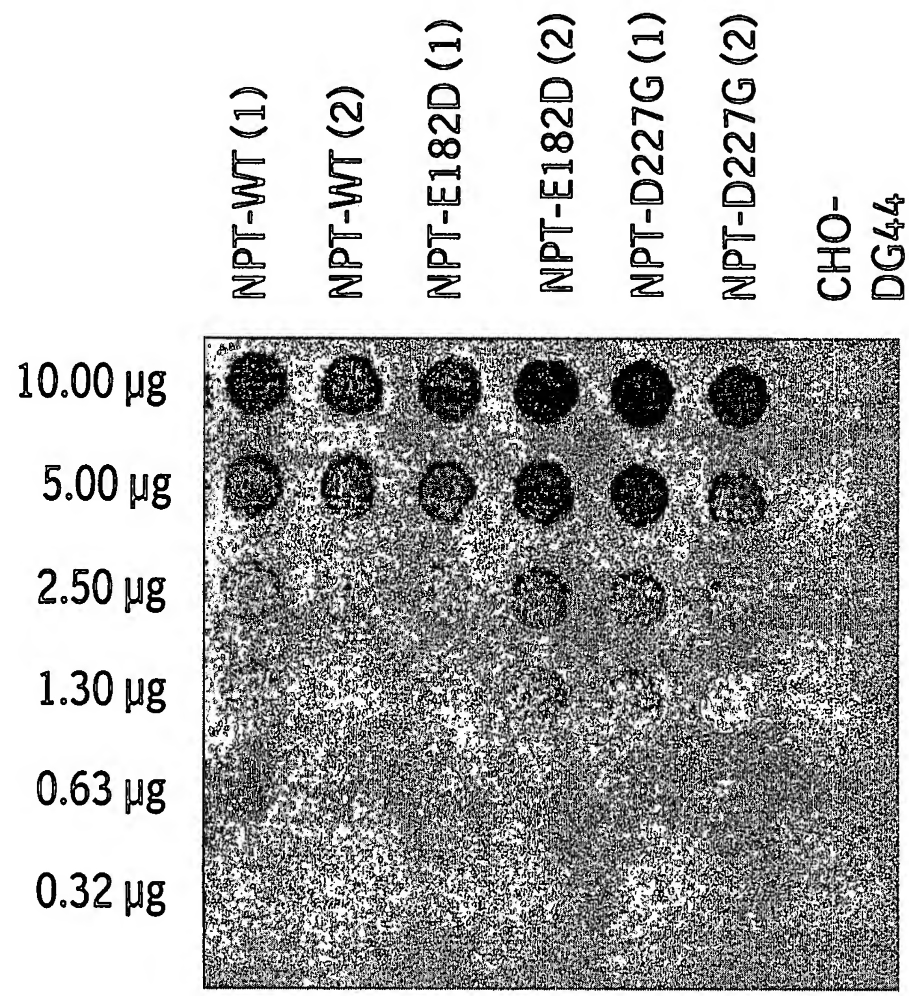

NPT-WT (1)
NPT-WT (2)
NPT-E182D (1)
NPT-E182D (2)
NPT-D227G (1)
NPT-D227G (2)
CHO-DG44
10.00 µg
5.00 µg
2.50 µg
1.30 µg
0.63 µg
0.32 µg

Figure 11

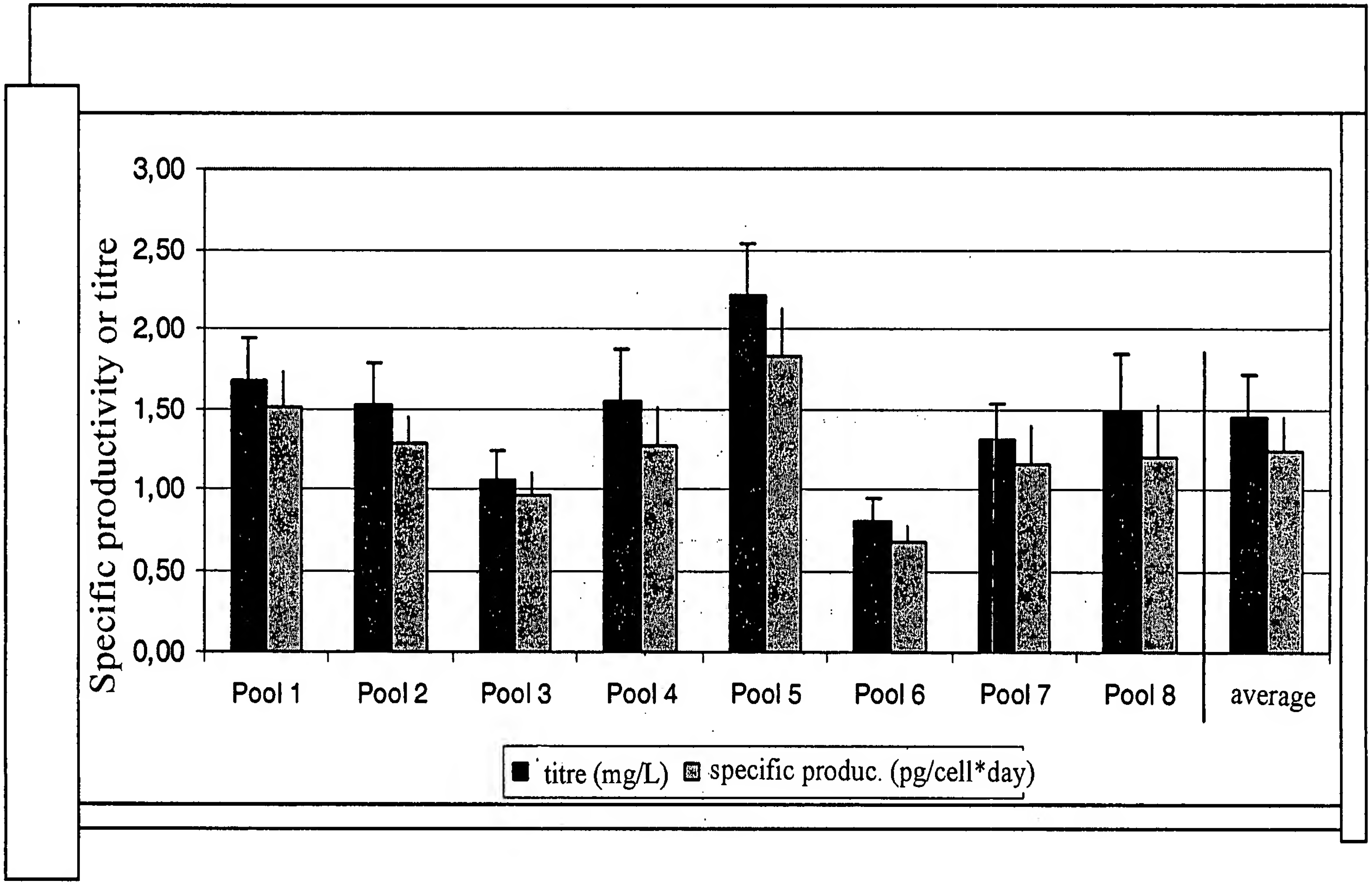

Specific productivity or titre
3,00
2,50
2,00
1,50
1,00
0,50
0,00
Pool 1
Pool 2
Pool 3
Pool 4
Pool 5
Pool 6
Pool 7
Pool 8
average
titre (mg/L)
specific produc. (pg/cell*day)

Figure 12

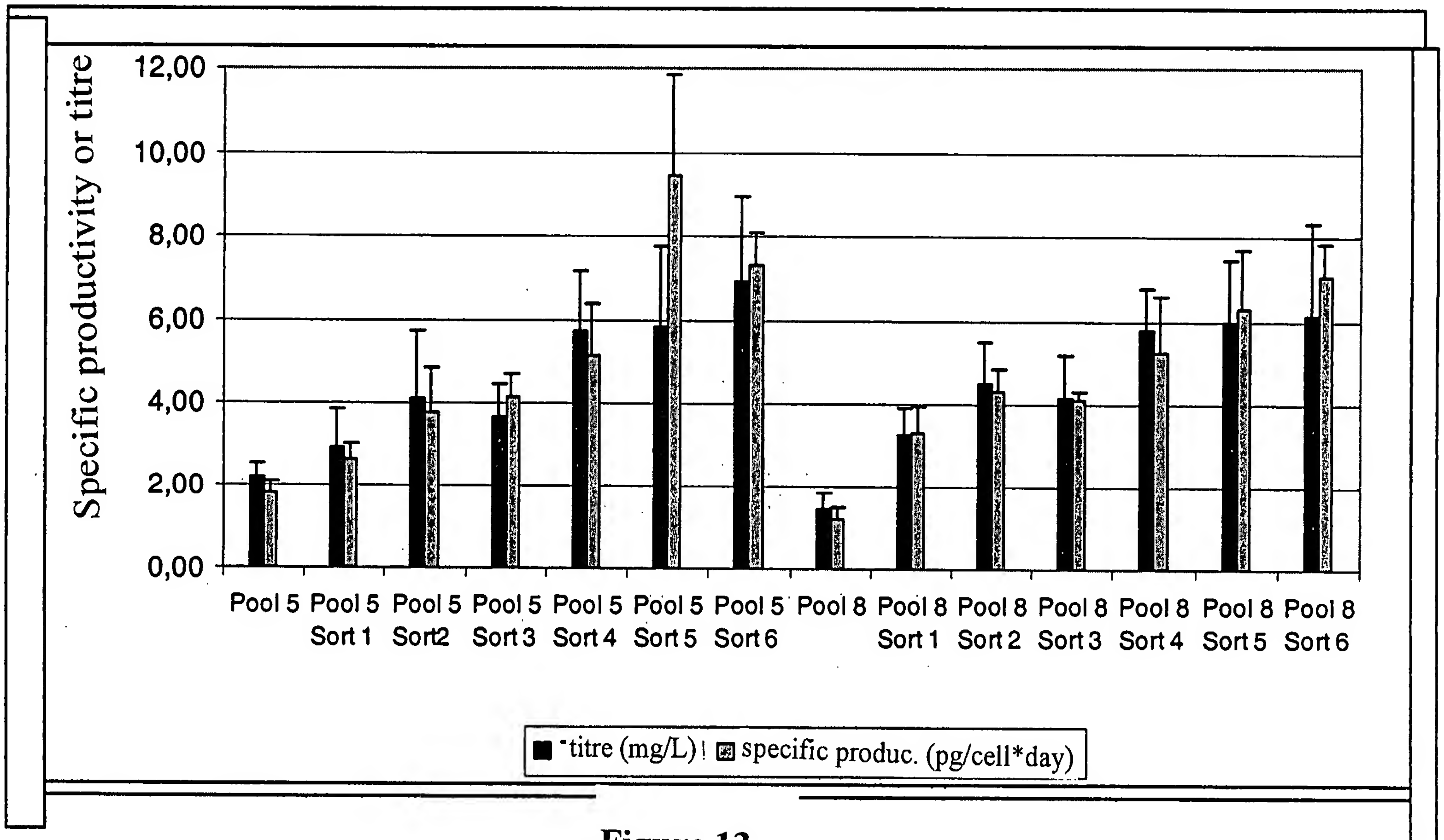

Specific productivity or titre
12,00
10,00
8,00
6,00
4,00
2,00
0,00
Pool 5
Pool 5 Sort 1
Pool 5 Sort2
Pool 5 Sort 3
Pool 5 Sort 4
Pool 5 Sort 5
Pool 5 Sort 6
Pool 8
Pool 8 Sort 1
Pool 8 Sort 2
Pool 8 Sort 3
Pool 8 Sort 4
Pool 8 Sort 5
Pool 8 Sort 6
titre (mg/L)
specific produc. (pg/cell*day)

Figure 13

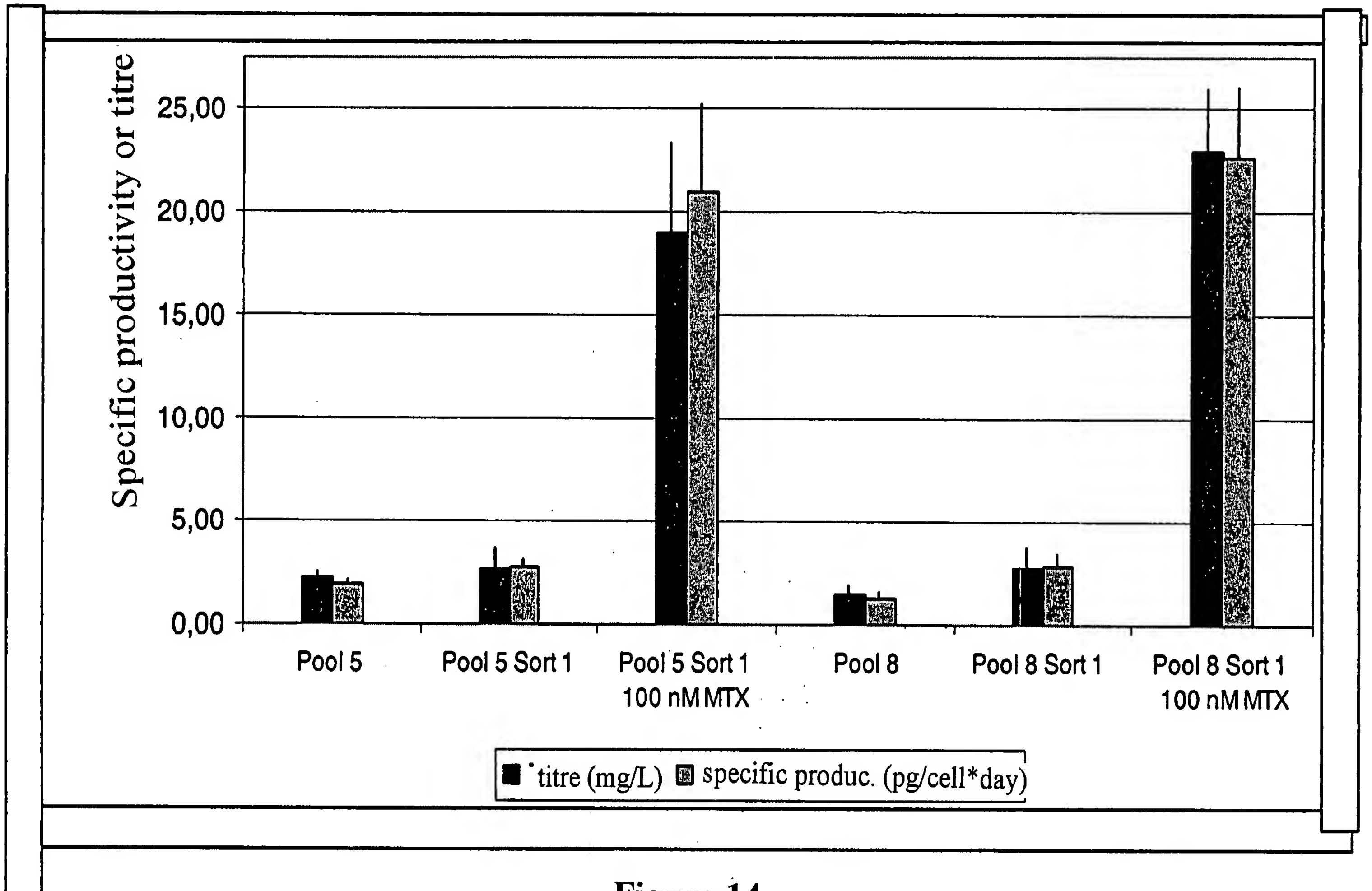
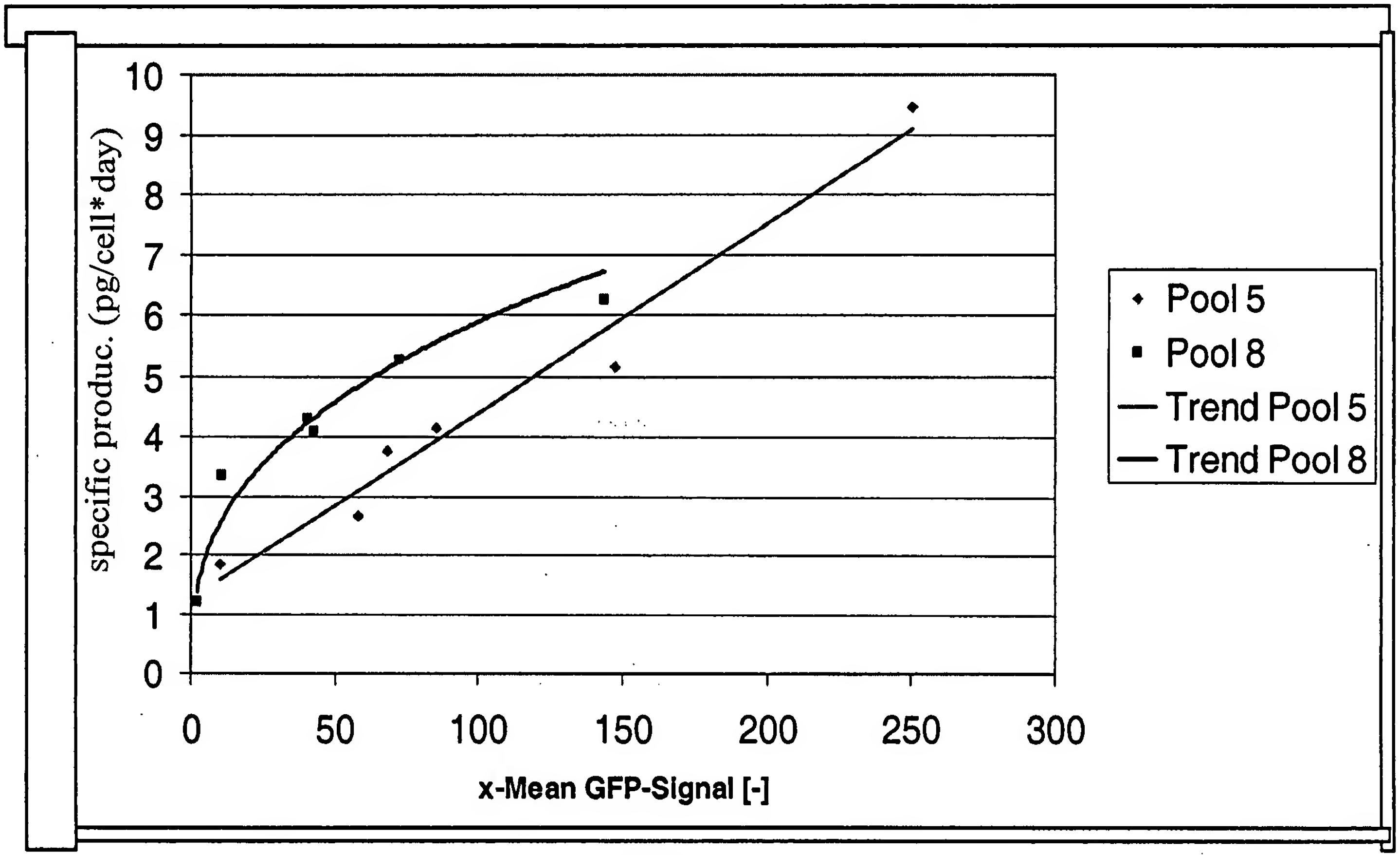

Specific productivity or titre
25,00
20,00
15,00
10,00
5,00
0,00
Pool 5
Pool 5 Sort 1
Pool 5 Sort 1 100 nM MTX
Pool 8
Pool 8 Sort 1
Pool 8 Sort 1 100 nM MTX
titre (mg/L)
specific produc. (pg/cell*day)

Figure 14



10
9
8
7
6
5
4
3
2
1
0
specific produc. (pg/cell*day)
0
50
100
150
200
250
300
x-Mean GFP-Signal [-]
Pool 5
Pool 8
Trend Pool 5
Trend Pool 8

Figure 15

ABSTRACT

Eukaryotic expression vector comprising a modified neomycin-phosphotransferase-gene, and methods for the selection of recombinant cells producing high levels of a desired gene product

The invention relates to a method for selecting highly productive recombinant cells, a method for preparing heterologous gene products and expression vectors which contain a modified neomycin-phosphotransferase-gene and a gene of interest functionally linked to a heterologous promoter and host cells transfected therewith which may be used in these processes.